Capacity-Approaching Modulation Formats for Optical Transmission Systems: Signal shaping and advanced de/muxing for efficient resource exploitation.

Estaran Tolosa, Jose Manuel

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Capacity-Approaching Modulation Formats for Optical Transmission Systems: Signal shaping and advanced de/muxing for efficient resource exploitation

Ph.D. Thesis

José Manuel Estarán Tolosa

November 1st, 2015

DTU Fotonik
Department of Photonics Engineering

DTU Fotonik
Department of Photonics Engineering
Technical University of Denmark
Ørsteds Plads 343
DK-2800 Kgs. Lyngby
Denmark

Preface

The work presented in this Thesis was carried out as a part of my Ph.D. project in the period September 30th, 2012, to November 1st, 2015. The work was conducted at various institutions including DTU Fotonik (Technical University of Denmark, Department of Photonics Engineering), Huawei Technologies Co. Ltd., Shenzhen, China, and Alcatel-Lucent Bell Labs, Villarceaux, France.

The Ph.D. project was partly financed by the VILLUM FOUNDATION, Søborg, Denmark, within the Young Investigator Programme and supervised by

- Darko Zibar (main supervisor), Associate Professor, DTU Fotonik, Technical University of Denmark, Kgs. Lyngby, Denmark
- Idelfonso Tafur Monroy (co-supervisor), Professor, DTU Fotonik, Technical University of Denmark, Kgs. Lyngby, Denmark
- Christophe Peucheret (co-supervisor), Professor, FOTON Laboratory (CNRS UMR 6082) - ENSSAT - University of Rennes 1, Lannion, France

Abstract

Aiming for efficient fiber-optic data transport, this thesis addresses three scenario-specific modulation and/or multiplexing techniques which, leveraging digital signal processing, can further exploit the available resources. The considered environments are: (i) (ultra) long-haul networks, where we focus on improving the receiver sensitivity; (ii) metropolitan area networks, where the target is providing spectral and rate adaptability with fine granularity and easy reconfigurability; and (iii) short-haul networks, where facilitating more affordable throughput scaling is pursued.

Functioning essentials, signal attributes, and digital processing structures are discussed in detail; whereby we give grounds for their adoption into the respective contexts of the three proposed systems. Eventually, we conduct the corresponding experimental validations; overall proving their potential. The reported results encourage the continuation of the research, mainly towards algorithm refinement.

Resumé

Denne afhandlings mål er at præsentere effektive fiber-optiske data transport og tre scenerier adresseres – mere præcist modulation og/eller data sammenfletnings teknikker som, ved hjælp af digital signal processer, kan udnytte de givne ressourcer bedre. Følgende scenarier bliver betragtet. (i) (ultra) lang distance netværk, hvor fokus er en forbedring af modtagerfølsomheden; (ii) metropol netværk, hvor målet er at forberede tilpasningen af både spektrum og hastighed med en bedre opløsning kombineret med nemme konfigurations muligheder ; (iii) kort distance netværk, med ønsket om at understøtte en mere økonomisk skalering af den samlede trafik.

De essentielle netværks funktioner, signal egenskaber samt digital signal proces strukturer bliver behandlet i detaljer, hvorefter der er basis for at tilpasse dem til den respektive kontekst i form af de tre foreslåede scenarier. Afslutningsvis bliver de tilsvarende eksperimentelle evalueringer udført, der samlet set understreger deres potentiale. De præsenterede resultater opfordrer til en forsættelse af forskningen indenfor dette emne med primær fokus på algoritme forbedringer.

Ph.D. Publications

The following publications have resulted from this Ph.D. project.

Articles in international peer-reviewed journals: (8)

J1 D. Zibar, L. H. H. Carvalho, **J.-M. Estarán Tolosa**, E. Porto Da Silva, C. Franciscangelis, V. Ribeiro, R. Borkowski, O. Winther, M. N. Schmidt, J. F. R. de Oliveira and I. Tafur Monroy, "Joint Iterative Carrier Synchronization and Signal Detection Employing Expectation Maximization," in *Journal of Lightwave Technology*, vol. 32, no. 8, pp. 1608-1615, April 2014.

J2 **J.-M. Estarán Tolosa**, D. Zibar, I. Tafur Monroy, "Capacity-Approaching Superposition Coding for Optical Fiber Links," in *Journal of Lightwave Technology*, vol. 32, no. 17, pp. 2960-2972, Sept 2014.

J3 M. Piels, E. Porto Da Silva, **J.-M. Estarán Tolosa**, R. Borkowski, I. Tafur Monroy, and D. Zibar, "Focusing over Optical Fiber Using Time Reversal," in *IEEE Photonics Technology Letters*, vol. 27, no. 6, pp. 631-634, March 2015.

J4 **J.-M. Estarán Tolosa**, M. A. Usuga Castaneda, E. Porto Da Silva, M. Piels, M. Iglesias Olmedo, D. Zibar, and I. Tafur Monroy, "Quaternary Polarization-Multiplexed Subsystem for High-Capacity IM/DD Optical Data Links," in *Journal of Lightwave Technology*, vol. 33, no. 7, pp. 1408-1416, April 2015 (invited).

J5 D. Zibar, L. H. H. Carvalho, M. Piels, A. Doberstein, J. Diniz, B. Nebendahl, C. Franciscangelis, **J.-M. Estarán Tolosa**, H. Haisch, N. G. Gonzalez, J. F. R. de Oliveira, and I. Tafur Monroy, "Application of Machine Learning Techniques for Amplitude and Phase Noise

Characterization," in *Journal of Lightwave Technology*, vol. 33, no. 7, pp. 1333-1343, April 2015.

J6 S. Saldaña Cercos, M. Piels, **J.-M. Estarán Tolosa**, M. A. Usuga Castaneda, E. Porto Da Silva, A. Manolova Fagertun, and I. Tafur Monroy, "100 Gbps IM/DD links using quad-polarization: Performance, complexity, and power dissipation," in *Optics Express*, vol. 23, no. 15, pp. 19954-19968, July 2015.

J7 A. Tatarczak, M. Iglesias Olmedo, T. Zuo, **J.-M. Estarán Tolosa**, J. B. Jensen, X. Xu, and I. Tafur Monroy, "Enabling 4-Lane Based 400â€‰G Client-Side Transmission Links with MultiCAP Modulation Advances in Optical Technologies," in *Advances in Optical Technologies*, vol. 2015, ID:935309, July 2015.

J8 M. A. Mestre, **J.-M. Estarán Tolosa**, P. Jennevé, H. Mardoyan, I. Tafur Monroy, D. Zibar and S. Bigo, "Novel coherent optical OFDM-based transponder for optical slot switched networks," in *Journal of Lightwave Technology* (invited).

Contributions to international peer-reviewed conferences: (14)

C1 **J.-M. Estarán Tolosa**, D. Zibar, A. Caballero, C. Peucheret, and I. Tafur Monroy, "Experimental Demonstration of Capacity-Achieving Phase-Shifted Superposition Modulation," in *European Conference on Optical Communication and Exhibition 2013*, ECOC 2013, paper We.4.D.5.

C2 D. Zibar, L. H. H. Carvalho, **J.-M. Estarán Tolosa**, E. Porto Da Silva, C. Franciscangelis, V. Ribeiro, R. Borkowski, J. F. R. de Oliveira, and I. Tafur Monroy, "Joint Iterative Carrier Synchronization and Signal Detection for Dual Carrier 448 Gb/s PDM 16-QAM," in *European Conference on Optical Communication and Exhibition 2013*, ECOC 2013, poster P.3.18.

C3 Y. An, M. Muller, **J.-M. Estarán Tolosa**, S. Spiga, F. Da Ros, C. Peucheret, and M.-C. Amann, "Signal Quality Enhancement of Directly- Modulated VCSELs Using a Micro-Ring Resonator Transfer Function," in *OptoElectronics and Communications Conference 2013* held jointly with *International Conference on Photonics in Switching 2013*, OECC/PS 2013, paper ThK3-3.

C4 **J.-M. Estarán Tolosa**, M. Iglesias Olmedo, D. Zibar, X. Xu, and I. Tafur Monroy, "First Experimental Demonstration of Coherent CAP for 300-Gb/s Metropolitan Optical Networks," in *Optical Fiber Communication Conference and Exhibition 2014*, OFC/NFOEC 2014, paper Th3K.3.

C5 T. Zuo, A. Tatarczak, M. Iglesias Olmedo, **J.-M. Estarán Tolosa**, J. B. Jensen, Q. Zhong, X. Xu, I. Tafur Monroy, "O-band 400 Gbit/s client side optical transmission link," in *Optical Fiber Communication Conference and Exhibition 2014*, OFC/NFOEC 2014, paper M2E.4.

C6 M. Iglesias Olmedo, A. Tatarczak, T. Zuo, **J.-M. Estarán Tolosa**, X. Xu, and I. Tafur Monroy, "Towards 100 Gbps over 100 m MMF using a 850 nm VCSEL," in *Optical Fiber Communication Conference and Exhibition 2014*, OFC/NFOEC 2014, paper M2E.5.

C7 **J.-M. Estarán Tolosa**, M. A. Usuga Castaneda, E. Porto Da Silva, M. Piels, M. Iglesias Olmedo, I. Tafur Monroy, "Quad-Polarization Transmission for High-Capacity IM/DD Links," in *European Conference on Optical Communication and Exhibition 2014*, ECOC 2014, paper PD.4.3 (postdeadline paper).

C8 D. Zibar, L. H. H. Carvalho, M. Piels, A. Doberstein, J. Diniz, B. Nebendahl, C. Franciscangelis, **J.-M. Estarán Tolosa**, H. Haisch, N. G. Gonzalez, J. F. R. de Oliveira, and I. Tafur Monroy, "Bayesian Filtering for Phase Noise Characterization and Carrier Synchronization of up to 192 Gb/s PDM 64-QAM," in *European Conference on Optical Communication and Exhibition 2014*, ECOC 2014, paper Tu.1.3.1.

C9 M. Piels, E. Porto Da Silva, **J.-M. Estarán Tolosa**, R. Borkowski, D. Zibar, I. Tafur Monroy, "DSP-Based Focusing over Optical Fiber Using Time Reversal," in *European Conference on Optical Communication and Exhibition 2014*, ECOC 2014, poster P.7.8.

C10 A. Tatarczak, Y. Zheng, G. A. Rodes, **J.-M. Estarán Tolosa**, C.-H. Lin, A. V. Barve, R. Honore, N. Larsen, L. A. Coldren, and I. Tafur Monroy, "30 Gbps bottom-emitting 1060 nm VCSEL," in *European Conference on Optical Communication and Exhibition 2014*, ECOC 2014, poster P.2.3.

C11 **J.-M. Estarán Tolosa**, D. Zibar, I. Tafur Monroy, "Capacity and Shaping in Coherent Fiber-Optic Links," in IEEE Photonics Conference 2014, IPC 2014, paper TuG2.4 (invited).

C12 A. Tatarczak, **J.-M. Estarán Tolosa**, M. Iglesias Olmedo, J. B. Jensen, J. J. Vegas Olmos, and I. Tafur Monroy, "Advanced digital signal processing for high-speed access networks," in *SPIE Photonics West 2015*, SPIE PW 2015, (invited).

C13 **J.-M. Estarán Tolosa**, X. Lu, D. Zibar, and Tafur Monroy, "Stokes Space in Direct-Detection Data Transmission Systems," in *Asia Communications and Photonics Conference 2015*, ACP 2015, paper AS4D.1 (invited).

C14 **J.-M. Estarán Tolosa**, M. A. Mestre, P. Jennevé, H. Mardoyan, I. Tafur Monroy, D. Zibar and S. Bigo, "Coherent Optical Orthogonal Frequency-Division Multiplexing for Optical Slot Switched Intra-Datacenters Networks," in *European Conference on Optical Communication and Exhibition 2015*, ECOC 2015, paper Tu.1.2.3.

Contents

Chapter 1

Introduction

Ever since the first steps back in 60s [1], fiber-optic data communications have experienced thriving evolution. Along these years, the emerging capabilities have encouraged incessant development of new applications and services with increasingly exigent technical requirements, in turn feeding back the need for progress at all levels; that is, network management and topology, degree of coverage and connectivity, and of course, achievable throughput. Notwithstanding, as this technology gets relentlessly pushed to the limits of performance, and the available resources to the edge of exhaustion, the evolution rate inescapably saturates. Broad-sense efficiency has then become a figure of radical importance for the current research force; such that, in the last years, it could be argued that a work paradigm has been widely adopted whereby not just *more* is encouraged but also *better used*. This is the big picture this investigation fits into.

1.1 General motivation

In this quest towards more efficient communications, the introduction of cost-effective high-speed digital signal processing (DSP) fostered a remarkable changeover around 00s. Advanced modulation formats and coding plus electronic compensation of linear and non-linear impairments became available, allowing for the simultaneous improvement of bandwidth use (spectral efficiency) and receiver sensitivity (power efficiency) with unmatched flexibility. This features naturally led to countless scenario-specific opportunities to the extent that, from dual-polarization long-haul transmissions to short-haul uncoded on-off-keying, transceivers implementing DSP are presently found everywhere.

The technical requirements on these transceivers, however, keep evolving with society's needs; and with them, the disquieting hurry for adaptation. In this direction, the global purpose of this Ph.D. investigation is: *by leveraging the ubiquity and potential of DSP, elaborating solutions going beyond existing ones in better exploiting the capacity of the medium and the available resources in a scenario-specific manner.*

1.2 Scope and outline

This thesis deals with the experimental validation of three concrete modulation and/or multiplexing techniques plus their associated transmitter/receiver algorithm structures. Aiming for more efficient communication, these schemes are each devised or adopted based on the specific techno-economical constraints and governing physics of their respective network segments:

- (Ultra) long-haul (Chapter 2) : superposition coded modulation with phase-shifted mapping [2–9] is proposed and experimentally demonstrated for the first time in coherent optical communications. This modulation scheme generates signals exhibiting two-dimensional Gaussian-like power distribution with lower processing complexity in both transmitter and receiver sides than similar alternatives. This motivates its consideration and for improving the receiver sensitivity of (ultra) long-haul lumped-amplified uncompensated links, where the non-linear interference is accurately modeled as additive Gaussian [10, 11].

- Metropolitan area (Chapter 3) : coherent multiple-band carrierless amplitude/phase modulation is introduced and experimentally demonstrated for the first time in coherent optical communications. In this modulation and multiplexing scheme the spectrum is digitally sub-divided into a user-defined number of multi-dimensional pass-band signals. High control of the spectral shape and line rate is accomplished through the variation of the center frequency, bandwidth, modulation format and dimensionality of each individual band, providing robustness against filtering effects without compromising spectral efficiency. These factors become particularly beneficial in metropolitan area networks, where the fast, diverse, and high-volume traffic increase [12] is progressively favoring the migration to high-speed all-optical reconfigurable mesh architectures [13–16].

- Short-haul (Chapter 4) : a technique enabling up to fourth-order polarization-multiplexing is conceived and experimentally demonstrated for the first time in an optical transmission system employing intensity modulation and direct detection. The proposed digital receiver operating on Stokes parameters performs blind and quasi-continuous tracking of characteristic Mueller channel matrix; hence supporting the simultaneous transmission of $\leq$4 states of polarization carrying independent data each. The use of complementary dimensions to wavelength and space in order to enhance short-haul networks is steadily more justified as affordable scalability gets more and more compromised with the increasing throughput demand [17].

These three self-contained chapters are all structured according to, first, the introduction of the scenario-specific motivation and problem statement; second, the contextualization of the work with respect to most-recent research activities plus explicit statement of our contribution; third, elaboration on the fundamentals that sustain the employed technique, as well as signal properties and major implementation concerns from both transmitter and receiver perspectives; fourth, explanation of the experimental goals, testbed, and discussion on the results; and fifth, synopsis, where the main points and findings are highlighted and possible future work lines proposed.

In the last block, Chapter 5, the whole investigation is summarized.

Chapter 2

(Ultra) long-haul networks

Carrying multi-terabit data traffic over reaches in the order of thousands of kilometers, these big "pipes" represent the spine of the global data communication network. Technology-wise, long-haul (LH) and ultra long-haul (ULH) optical data links employ wavelength-division multiplexing (WDM) in various configurations. The use of lower-rate channels reduces the cost per transmitted bit, which comes in addition to the functional throughput scaling capabilities that WDM inherently features. For instance, assuming 50-GHz inter-channel spacing[1], up to 160 data-independent channels may be multiplexed across the entire C and L bands. As far as transceivers are concerned, most-recent implementation agreements define 100-Gbps coherent technology based on polarization-multiplexed quadrature phase-shift keying (PM-QPSK) with 2 bit/s/Hz of spectral efficiency (SE) [19]. Several technological alternatives for next-gen 400-Gbps coherent transceivers are elaborated in detail in [13]. Channel coding blocks typically employ hard-decision 7%-overhead Reed-Solomon forward error correction (FEC) [20]. More sophisticated recommendations with improved performance are described by the ITU-T under *super-FEC* denomination, covering mixed alternatives with 7%-, 11%- and 25%-overhead as well as hard- and soft-decision approaches [20]. Software-configurable FEC options are also supported in some commercially available cards.

These networks describe centralized point-to-point or few-node distributed mesh topologies in which assuring light-path integrity[2] and sig-

[1]Nominal inter-channel spacing as recommended by telecommunication standardization sector of ITU (ITU-T) include 12.5, 25, 50 and 100 GHz. ITU-T also defines the nominal central frequencies for each grid [18].

[2]Virtue of the light-path to remain operative.

nal integrity[3] becomes critical. Whereas the former requires standalone improvements in the reliability of the link's components (e.g. fibers, cables, optical amplifiers or optical add-drop multiplexers), enhancing the signal integrity demands measures essentially opposed to those for increasing throughput (e.g. increasing inter-channel spacing or reducing launch power to diminish nonlinearities). The latter hints one conclusion of practical importance [21]: the existence of an optimum operation point for each system configuration which establishes the throughput-times-distance product performance beyond which reliable transmission is not possible. Faced with need for longer distances and/or higher throughput [12] as well as economical restrictions that impede riotous fiber deployment [22], the scrutiny of the boundaries of achievable capacity has gained great attention [23–30].

The optimum operation point results from trying to overcome the optical bandwidth scarcity and/or insufficient values of optical signal-to-noise ratio (OSNR). This premise evinces the appropriateness of maximizing both the SE and the receiver sensitivity, which directly relates to the optimization of the inter-channel spacing, the per-channel bandwidth, and the per-channel modulation format plus de/coding scheme. Migrating from coarse frequency grids, highly spectral-efficient and flexible solutions[4] are attained thanks to tighter channel bandwidth assignation and fine guardband reduction [32, 33], ultimately limited by detrimental effects like spectral clipping and inter-carrier overlapping. It is precisely in the limit of channel packing density that advanced modulation formats (AMF) become major drivers of the SE and the receiver sensitivity for a given transmission reach. Richer constellations provide greater information rates per Hz at the expense of signal power and noise tolerance reduction [34], which may not be acceptable. Interestingly, inspired by the progress in the design of application-specific integrated circuit (ASIC) and digital-analog interfaces, DSP has allowed for the relaxation of such compromise, bringing forward the *processing complexity* as additional resource to trade on [35].

With coherent detection enabling digital manipulation of the optical field, uncompensated polarization-multiplexed transmission over long distances signals is a reality. Continuous algorithm refinement has per-

[3]Virtue of the signal to meet the quality criteria that enables successful data transmission and recovery.

[4]Slot granularities of 12.5 GHz are being studied. This fragmentation allows for fitting 100 Gbps within 37.5 GHz instead of the standard 50 GHz window (33% spectral savings) [31]. Anticipating future needs, other research activities contemplate 6.25-GHz slot granularity (see footnote 1).

mitted close to total compensation of linear distortions (e.g. chromatic dispersion (CD) and polarization-mode dispersion (PMD)) to an extent that, at present, non-linear signal-noise interaction totally governs the engineering of the optimum throughput-times-distance product [36]. This outlook has fostered research activities towards exploring the underlying physics for accurate channel modeling, paving the way for the development of manifold DSP methods aiming nonlinearity compensation. Some of the most explored algorithms for optical communications include back-propagation [37] and variants[5], inverse Volterra series [43], phase-conjugated twin-waves [44], mid-span optical phase conjugation [45], frequency pilot tone conjugation [46] or non-linear Fourier transformation [47]. Unfortunately, despite the efforts to improve the trade-off between implementation complexity and performance, achievable sensitivity gains have proved insufficient[6] within acceptable cost and power consumption margins. As a result, techniques for nonlinearity mitigation (i.e. nonlinearity minimization during propagation [49, 50] and/or nonlinearity tolerance maximization [51, 52]) are being considered to further boost the receiver sensitivity; and in this direction, coded modulation combined with constellation shaping [53, 54] has the potential to achieve constrained capacity by itself [55, 56].

Even though shaping the power distribution of the transmitted signal can extend the achievable throughput to its upperbound [55, 57], practical restrictions (ranging from technological limitations to fundamental non-linear noise understanding) have forced much humbler aspirations as of today [23, 49, 50]. Nevertheless, the undisputed potential of this technique has motivated the study and perfection of existing candidates towards their adaptation to different scenarios in fiber-optics data communications (see Section 2.1). Among those, certain variations of superposition coded modulation (SCM) [7, 58–61] have been proved to approach continuous memoryless additive white Gaussian noise (AWGN) channel capacity (see Appendix B) with low processing complexity, while providing convenient modulation order adaptability and simplified demapping. Resorting to the Gaussian noise model [29], the latter approach arises as a promising candidate for application in the scenarios covered in this chapter. Originally suggested by H. Schoeneich and P. A. Hoeher for wireless communications [2]

[5]Some examples include stochastic [38], folded [39], weighted [40], perturbation [41] and filtered [42] back-propagation.

[6]in the sub-decibel range in typical densely populated WDM LH/ULH scenarios, employing uncompensated lumped-amplified standard single mode fiber (SSMF) transmission links and free-running lasers [48]

and successively researched [3–9], we propose and experimentally demonstrate for the first time the use of SCM with phase-shifted mapping (PSM) in optical communications.

This chapter is subdivided into four blocks. First, this investigation is positioned with respect to the latest research, narrowing down the scope to coded modulation schemes employing Gaussian shaping for application to optical data links where the Gaussian noise model holds strictly. Second, the properties of PSM signals are elaborated together with the essentials concerning digital generation, detection and decoding. Third, the experimental setup and results are presented and discussed, focusing on the implementation feasibility when standard DSP algorithms are employed for signal recovery. Finally, the work is summarized, comprehending a motivation recap, main results and conclusions, and plausible future work lines.

2.1 Work positioning

Two aspects shall be used to classify diverse constellation shaping techniques: (i) the optimization criterion, and (ii) the statistical properties of the signal impairments, in turn subject to the particular link components and transmission technology. In this chapter, the focus is on mutual information (MI) maximization for coherent LH/ULH uncompensated lumped-amplified SSMF transmission links employing WDM. Customary treated as noise, non-linear interference in these scenarios is accurately modeled as additive Gaussian [10, 11], for which the optimum source distribution in MI sense has been known with exactitude for more than half a century [62]. This fact has motivated the application of mature constellation shaping techniques to current fiber-optic communications and the immediate extrapolation of certain results for performance estimation and capacity lower-bounding.

In the following, we present the most relevant investigations on MI-wise constellation shaping for the above-mentioned scenarios, subdivided into probabilistic and geometric methods. Our work is positioned with respect to the literature and major contributions highlighted.

Probabilistic shaping

In this family of techniques, the power distribution across the constellation space is altered by varying the individual probability of occurrence of

the symbols through the inducement of controlled redundancy or bit-to-symbol mapping ambiguities. The constellation grid remains unmodified. Simulation and/or theoretical results include:

- [63]: adaptation of Trellis shaping for attaining Gaussian distribution and high-speed efficient decoding.
- [64]: four-dimensional rate-adaptive coded modulation with shaping capabilities based on low-density parity-check (LDPC) codes and Shell mapping.
- [65]: arithmetic distribution matching in combination with systematic LDPC codes for enabling bit-level shaping and decoding [66]. No iterative decoding is needed. Maximum reach extensions of 8% and 15% are shown for 16-ary and 64-ary shaped quadrature amplitude modulation (QAM) respectively.
- [67]: modification of the pragmatic shaping algorithm proposed by D. Raphaeli which outperforms the original for high signal-to-noise ratio (SNR). Theoretical SE gains around 0.35 bit/s/Hz are reported in the optimum operation point for shaped 1024-ary QAM.

To the best of our knowledge, only one experimental demonstration of probabilistic shaping has been reported in coherent optical communications:

- [68]: employing arithmetic distribution matching, the authors demonstrate shaped and rate-adaptive transmission with fixed channel-code overhead. This is possible owed to the bit-level redundancy that ADM process introduces, then diminishing the effective transmission rate without having to modify the channel code. The results show showing 15% increase in the achievable throughput and 43% increase in the achievable reach with respect to uniform 200-Gbps 16-ary QAM.

Geometric shaping

The power distribution across the constellation space is modified via direct reallocation of the conforming symbols while maintaining uniform probability of occurrence. Simulation and/or theoretical results include:

- [23]: constrained-capacity increase by optimizing the inter-ring spacing and the ring occupation density of uniform continuous ring constellations. Gain in SE in the order of 0.2 bit/s/Hz is obtained in the

corresponding point of optimum operation for shaped 2-ring constellation.

- [69]: a systematic method to find MI-wise optimized constellations in minimum mean square error (MMSE) sense is presented within the context of high-speed optical links.

- [70]: capacity-approaching (in memoryless AWGN channel) iterative polar modulation (IPM) is introduced, a modulation format obtained through the iterative quantization of a bivariate Gaussian distribution while minimizing the quantization mean-square error (QMSE).

The experimental demonstrations cited below need clarification. Although the employed link infrastructures do not fully correspond to the one of interest, the shaping optimization criterion and the assumptions on the transition probability density function (pdf) of the channel converge to the ones discussed. Because these investigations partly deal with general implementation concerns, their relevance in our context is justified:

- [71]: 256-ary IPM on orthogonal frequency-division multiplexing (OFDM) signal is transmitted over 800-km uncompensated raman-amplified ultra-large-area fiber (ULAF). Up to 6 dB of power budget margin increase compared to the same order QAM is shown.

- [72]: following up on [71], terabit transmission of OFDM signal over Korean deployed legacy WDM link employing 256-ary IPM is demonstrated. Enhancements in the soft-based decoding process are introduced.

Constellation shaping techniques incorporated within coded modulation schemes are meant to further increase the power budget[7], hence potentially coping with the performance penalty induced when extending the reach and/or increasing the symbol rate. Nonetheless, both longer reach and boosting the baudrate worsen the noise severity via linear and non-linear mechanisms, imperiling the sensitivity improvement and thus questioning the utility of the typically "very high complexity" [72] and hard-scalable shaping procedures. Because such gains are already diminished by the assumptions on the transition pdf of the channel, functional schemes worth consideration for implementation must accomplish non-negligible shaping

[7]Understood as the range of propagated optical power values for which, given predefined link characteristics, signal integrity is preserved (see footnote 3)

gains maintaining low-complexity generation and optimization. This is exactly the strength of SCM-PSM over other solutions presented to date, a transmission model that approaches memoryless AWGN channel capacity asymptotically with the number of bits per symbol, while offering enhanced simplicity and flexibility at the transmitter and receiver sides.

Being SCM-PSM the focus of this investigation, in the next sections we elaborate on the essentials and main properties of this modulation format, and present the results of the first experimental generation, transmission and detection in optical communications. Theoretical shaping gains with respect to uniform QAM are provided for memoryless AWGN channel, as well as the information rates for up to 12 bit/sym configuration. The system performance is experimentally investigated for dual-polarization (DP) single-carrier 16-ary, 32-ary and 64-ary PSM in optical back-to-back (B2B) configuration, and after 240-km uncompensated lumped-amplified SSMF link.

Contributions of this work

- Original idea: proposal of SCM-PSM for its use in uncompensated lumped-amplified SSMF transmission links employing WDM with the goal of increasing constrained capacity.
- Theoretical characterization: the exact values of information rate, shaping gain and absolute gap to capacity are calculated. Likewise, a detailed normality test including two-dimensional (2D) Gaussian fitting is realized.
- Experimental results: experimental demonstration of SCM-PSM in coherent optical communications.
- Related first-author publications: [73] (conference), [74] (journal) and [75] (invited conference).

2.2 Superposition Coded Modulation with Phase Shifted Mapping (SCM-PSM)

Constellation shaping in SCM-PSM is passive. The superposition of independent streams of weighted binary phase-shift keying (BPSK) symbols provides a simplified, flexible and effective means of fitting the capacity-achieving Gaussian distribution with monotonically increasing precision.

Non-bijective[8] constellations may occur for certain configurations, which together with the sub-optimal symbol allocation in symbol error rate (SER) sense, suggest the use of maximum a posteriori probability (MAP) iterative demodulation. Interestingly, the superposition-based generation allows for considerable simplification of the soft-input soft-output (SISO) demapper, thus conveniently alleviating the DSP complexity ratio between receiver (RX) and transmitter (TX). These aspects are detailed in the following subsections.

2.2.1 Generation and signal properties

Devised by L. Duang et al. in 1997 [58], the exploitation of the central limit theorem to create discrete constellations with arbitrarily accurate Gaussian power distribution is a mature concept. The passive generation of symbol distributions that asymptotically close the gap to the Shannon limit and the possibility of conducting algorithmic de-/mapping, motivated the study of this technique as part of sophisticated coded modulation schemes. In our case, we employ PSM-based shaping approach within the SCM framework with single-layer encoding strategy, bit-wise interleaving and iterative decoding-demapping (ID) (identical to bit-interleaved coded modulation (BICM) with ID).

In this subsection we spotlight the shaping mapper, i.e. PSM; a functional configuration that yields 2D constellations with near 1 average information rate per bit, yet effectively approaching capacity.

2.2.1.1 Digital transmitter

Rather than performing non-linear bit-to-symbol mapping, the output symbols are designed by linearly superimposing certain number of weighted parallel data streams. This readily scalable and optimizable structure, facilitates information rate adjustment and the arrangement of diverse power distributions by modifying the number of overlapped sequences and the multiplicative factors. Besides, the fact that each branch may contain independent data with different coding strategies suggests its application as medium access control (MAC) technique in multi-user scenarios. Mathematically, SCM with generalized mapping can be expressed as:

[8]Bijective modulation: the constellation cardinality equals the number of unique bit combinations for a fixed bit-tuple length (number of superimposed branches in SCM). Non-bijective modulation: the constellation cardinality is lower the number of unique bit combinations for a fixed bit-tuple length.

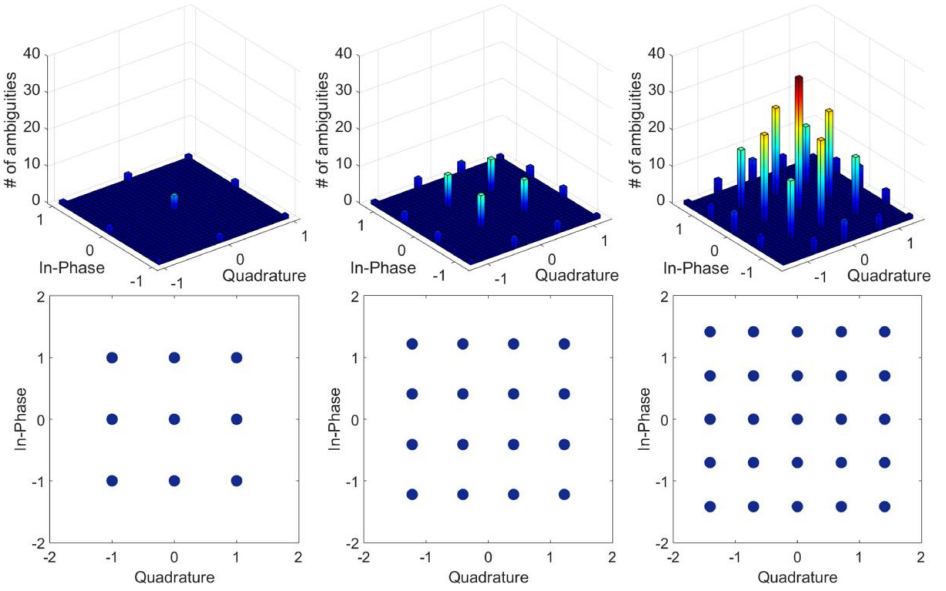
$$y = y_I + jy_Q = \sum_{n=1}^{N} e^{j\pi b_n} h_n \text{ with } h_n = \alpha_n e^{j\theta_n} \tag{2.1}$$

where y_I and y_Q are the in-phase and quadrature components, N establishes the number of superimposed branches (bits per symbol), b_n corresponds to the encoded binary bits and h_n represents the complex weights that determine the symbol mapping scheme. It is now apparent that α_n and θ_n constitute the mapping parameters, thus their concrete selection fully dictates the properties of the output constellation. For instance, it is possible to select between *probabilistic shaping* or *geometric shaping*. An example of the former is the *type-I sigma mapper* [59], which reaches theoretical memoryless AWGN channel capacity when $N \to \infty$, and whose parameters are given by:

$$\begin{aligned} &\alpha_n = 1 \text{ for } 1 \leq n \leq N \\ &\theta_n = \begin{cases} 0 & \text{for } 1 \leq n \leq \frac{N}{2} \\ \frac{\pi}{2} & \text{for } \frac{N}{2} \leq n \leq N \end{cases} \end{aligned} \tag{2.2}$$

The output constellations with *type-I sigma* for $N = 4, 6$ and 8 along with the corresponding 2D histogram are shown in Figure 2.1. Even though it is not surprising that the goodness of fit is inversely proportional to the number of ambiguities, the information rate increase is only logarithmic with N with *type-I sigma mapper* [3]; obviously slower than the number of bits per symbol, N. Resolving the numerous uncertainties for large N requires channel coding to induce equivalent redundancy[9] for reliable transmission (>50% for $N > 8$), thereof drastically decreasing the throughput under realistic implementation constraints (i.e. finite resources including affordable latency and implementation penalties). For example, one important design constrain in LH/ULH systems is per-carrier throughput maximization which, in case of severe information rate penalties, leads to technologically unsupportable constellation expansion ratio (CER). One intuitive solution to increase the cardinality for a given N is to assign different power coefficients (α_n) to all the branches so as to force bijectivity. This mapper is called *type-II sigma* [59] and belongs to the *geometric shaping* group.

[9]Shaping techniques based on the superposition of independent data streams are redundancy-free, thus the equivalent information rate reduction can be interpreted as the necessary channel coding-induced redundancy that enables reliable transmission.



(a) $N = 4$ (b) $N = 6$ (c) $N = 8$

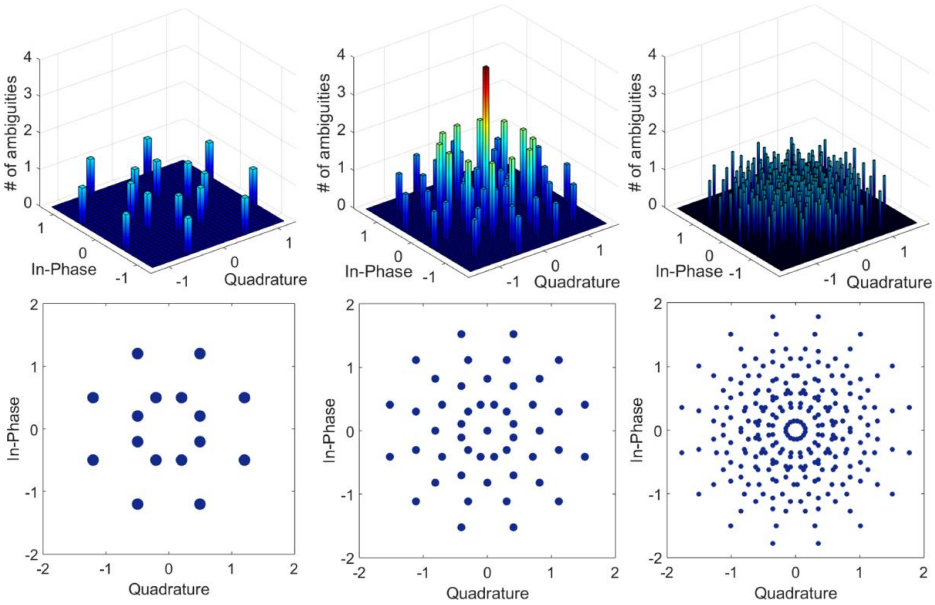
Figure 2.1: Constellation diagrams (bottom) and 2D histograms (top) of SCM with *type-I sigma mapper* for $N \in \{4, 6, 8\}$.

Unfortunately, the goodness of fit is strongly diminished, and the variable power scaling across the branches poses undesirable energy inefficiency.

In the case of PSM configuration, each branch is applied a different constant-amplitude complex weight. This simple procedure provides further symbol allocation diversity on the constellation diagram, increasing the cardinality and maintaining equal peak magnitude in all the parallel data streams. The latter turns convenient for analogue (digital-to-analogue converter (DAC)-free) implementation. The mapper parameters are given by:

$$\begin{aligned} \alpha_n &= 1 \text{ for } 1 \leq n \leq N \\ \theta_n &= \tfrac{\pi(n-1)}{N} \text{ for } 1 \leq n \leq N \end{aligned} \tag{2.3}$$

The output constellations with PSM for $N = 4, 6$ and 8 along with the corresponding 2D histograms are shown in Figure 2.2. As previously pointed out and inferred from the existence of ambiguities in the middle case, the inter-layer interference may give rise to non-bijective mapping for certain values of N, yielding a predominantly (but not totally) *geometric*



(a) $N = 4$ (b) $N = 6$ (c) $N = 8$

Figure 2.2: Constellation diagrams (bottom) and 2D histograms (top) of SCM with PSM for $N \in \{4, 6, 8\}$.

shaping with 90%<rate≤100% (see Subsection 2.2.1.2). Nevertheless, it is noteworthy that ambiguous mapping may be beneficial for some modulation orders, where the constellation's inter-symbol minimum Euclidean distance is increased with respect to the bijective equivalent for the same average power [7]. Consequently, the receiver sensitivity is nominally increased in addition to the shaping gain[10]. Assuming standard memoryless demapper,

[10]Because the contribution to the error rate of each constellation symbol is directly related to its probability of occurrence, such probabilities shall be taken into consideration for estimating the performance and for enabling fair comparison among different mapping approaches. One possible figure to coarsely observe the influence of those probabilities is the statistically averaged minimum Euclidean distance, in which a simple probability-dependent weighting of the per-symbol minimum Euclidean distance is applied. The expression is given by:

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \min(d(x_i, x_j)) p(x_i),$$

where x_i and x_j are complex values representing concrete symbol locations from the reference constellation, N is the constellation cardinality, $\min(\bullet)$ is the minimum opera-

inter-symbol minimum Euclidean distance is consistently related to SER in channels exhibiting additive circularly symmetric noise. Needless mentioning that symbol labeling optimization is crucial for assuring low-error binary data recovery, notably challenging task given irregular symbol grids.

2.2.1.2 Information rate

It is now clear that determinate mapping parameters in SCM produce ambiguous bit-to-symbol mapping that reduces the average entropy per bit, then directly influencing the information rate, the code design and the receiver structure. Understanding and quantifying the achievable rates turns of importance for the subsequent code design and thus determining what configurations suit the systems' requirements. In this subsection, the exact numbers for SCM-PSM are presented. Despite other authors have carried out similar exercise [3], our results introduce corrections in the accuracy and add extra cases of study.

Following general information theory one can develop SCM's theoretical information rate for any given mapping. Because ideal generation is noiseless, the modulation information rate results in:

$$I(X;Y) \triangleq H(Y) - H(Y|X) = H(Y) \tag{2.4}$$

where $I(X;Y)$ is the mutual information between the output random variable Y and input random variable X to the SCM structure, and $H(\bullet)$ is the discrete entropy operator. Now, applying 2.4 to SCM case:

$$I(y;\alpha_1 e^{j(\pi b_1+\theta_1)},\ldots,\alpha_n e^{j(\pi b_n+\theta_n)}) = H(y) \tag{2.5}$$

where $y \in y_1, y_2, \ldots, y_K$ represents the set of possible output complex values (K is the cardinality and it accounts for ambiguities caused by nonbijective mapping). Since the discrete entropy of the linear addition of the constituent antipodal binary symbols will always be equal to the number of conforming layers N, we can write:

$$H(y) = \sum_{m=1}^{K} P(y_m)\log_2 P(y_m) \leq -\sum_{n=1}^{N} H(\alpha_n e^{j(\pi b_n+\theta_n)}) = N \tag{2.6}$$

from which follows that the ideal transmitter output information rate is upper-bounded as:

tor, $d(\bullet)$ is the Euclidean distance operator and $p(x)$ is the probability of occurrence of symbol x.

$$Rate \leq \frac{H(y)}{N} \tag{2.7}$$

Inequality 2.7 makes explicit a well-known result: the achievable information rate is maximized when the output distribution from the mapper is statistically uniform given a fixed input bit-tuple length (N). In this regard, it should be understood that pure *geometric shaping* is a redundancy-free process that does not condition the channel code in the absence of noise. Although this case does not strictly apply to all PSM cases, the average rate never goes below 91.6% for up to 12 bits per symbol. Table 2.1 summarizes these and other results in a roster of the most relevant parameters of the output constellations, including bit-tuple length, cardinality, discrete output entropy and resulting information rate per bit.

Table 2.1: Roster of information theory data of PSM signaling. The constellation cardinality, its entropy and the average information rate per bit are listed for $\{N \in \mathbb{N}; 2 \leq N \leq 12\}$.

N	Cardinality	Entropy [bits]	Inf. rate per bit
2	4	2.000	1.000
3	7	2.750	0.916
4	16	4.000	1.000
5	31	4.937	0.987
6	49	5.500	0.916
7	127	6.984	0.997
8	256	8.000	1.000
9	343	8.250	0.916
10	961	9.875	0.987
11	2047	10.999	0.999
12	2401	11.000	0.916

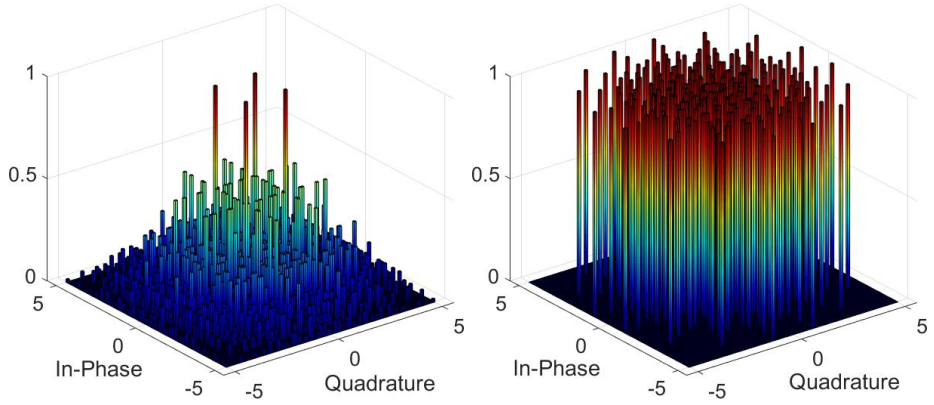
Variations in the information rate can also be qualitatively evaluated from the output constellation histograms (see Figure 2.2). Unlike $N = 4$ and 8, for which the probability of occurrence of all the symbols equals $\frac{1}{N}$, 64-ary PSM shows clear non-bijective behavior. The lowest power symbol (0 power in this case) is attained with up to 4 different combinations of 6 bits, and the symbols in the lowest-power middle ring are obtained with 2 unequal bit-tuples each. On this account, the total number of unique discrete power locations (cardinality) is reduced to 49, which yields one

of the lowest information rates in PSM; anyway imposing as low as 8.4% worst-case channel coding redundancy, well within the 25%-overhead FEC considered for high-speed LH/ULH coherent systems [20].

2.2.1.3 Gaussianity and shaping gain

Owed to diverse technological limitations (e.g. processing power cap and partial knowledge of the interfering spectrum), it is customary to resort to the average transition pdf of the optical channel for the calculation of pragmatic capacity lower-bounds and the associated constellation shapes. The harshness of this approximation is often relaxed assuming nonlinearity compensation on the channel of interest, removing thereby great part of the CD-induced memory and other intra-carrier effects. Interestingly, regardless of the use of compensation techniques, the time-averaged overall noise interference (linear and non-linear) in the scenarios of concern has been long proved additive Gaussian; for which the capacity-achieving input power distribution is known with exactitude: continuous zero-mean circularly symmetric normal. This entails two important observations, (i) Gaussian-like constellation shaping can improve the receiver sensitivity (shaping gain) both in the linear and non-linear (or mixed) regimes; and (ii) conditioned to their cautious interpretation, well-known capacity results for memoryless AWGN channel can be extrapolated to the concrete LH/ULH optical links discussed here, thus facilitating analogue analyses. Relying on these points to support the relevance of the forthcoming study, this subsection elaborates on the Gaussian-fitting accuracy of SCM-PSM (Part I), and its influence on the theoretical SE and shaping gain over QAM modulation (Part II) in the presence of memoryless AWGN.

Gaussianity : Any joint multi-dimensional distribution can be generated, and therefore studied, from the constituent marginals and the mathematical description of their dependence (*copula*). In our case the consideration framework is narrowed down to 2D output distributions, hence two random variables corresponded to orthogonal quadrature components. When general SCM is contemplated, the marginal distributions are given by the superposition of antipodal symbols, singly scaled according to predefined mapping rules that also determine the *copula*. It follows from the previous observation that not all the weight configurations enable the strict application of the central limit theorem (CLT) and/or guarantee statistical independence. For instance *type-I sigma mapper* (see 2.2) ensures the variables to be marginally Gaussian (classical CLT formulation holds) and



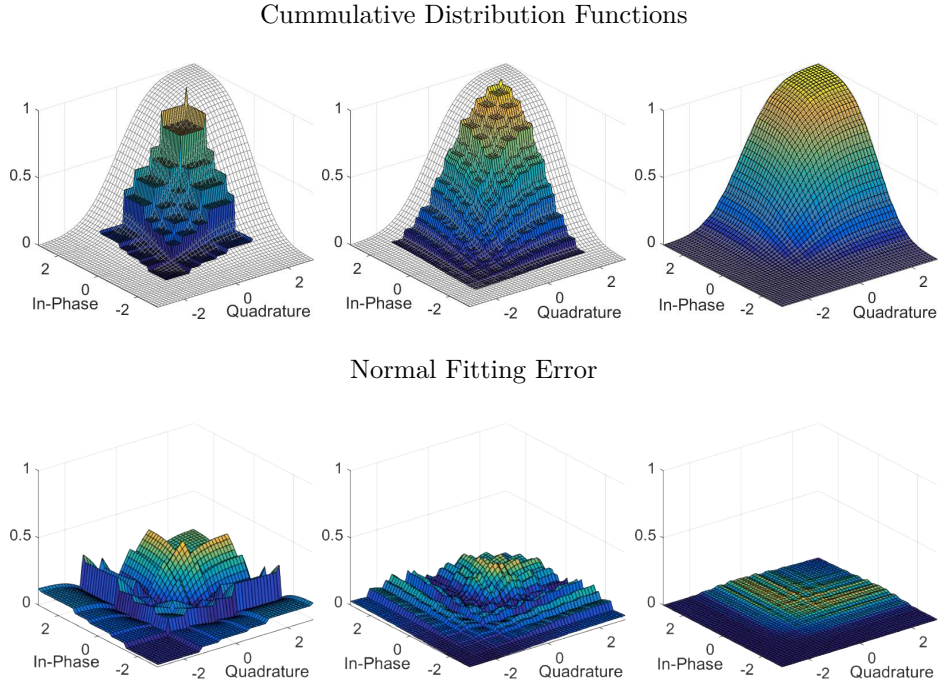
(a) Stat. indep. quadratures (b) Stat. dep. quadratures (PSM)

Figure 2.3: 2D pmf of the $N = 8$ output signal when the same marginal is uniformly sampled for each quadrature independently, and when those marginals are dependent according to PSM equations.

independent, hence trivially leading to circularly-symmetric[11] normal distribution (see evolution from Figure 2.1a to Figure 2.1c). In contrast, the PSM structure (see in 2.3) couples the quadrature components via non-orthogonal phase rotations across the branches (notice the uniform phase distribution between 0 and $\frac{\pi(N-1)}{N}$), whose unequal Cartesian projections prevent classical CLT from holding. Although the asymptotic Gaussianity of the marginals can be easily demonstrated (see appendix A), the proof of joint 2D-normal distribution is not so immediate, as the degree of dependence changes for different values of N. As illustrative evidence of the dependence between the quadrature components in SCM-PSM, Figure 2.3 shows the 2D pmf of the output distribution for $N = 8$ when identical marginals are independently sampled with a uniform distribution (2.3a), and when they are dependent according to PSM rules (2.3b). The observed differences confirm the dependency of the quadratures components in PSM.

In order to evaluate the point-wise convergence to bivariate Gaussian distribution of SCM-PSM constellations, a normality test is performed by comparing theoretical and empirical cdfs for increasing values of N. The metrics are both qualitative, through visual correlation of the overlapped

[11] The orthogonality of the quadrature vectors imply uncorrelation when their individual mean equals zero.



(a) $N = 4$ (b) $N = 6$ (c) $N = 12$

Figure 2.4: 2D cdf of PSM output distributions plus ideal unit-variance Gaussian distribution (top), and the corresponding fitting error (bottom) for $N \in \{4, 6, 12\}$.

cdfs and the areas of discrepancy, and quantitative, with the percentage of average fitting accuracy. Figure 2.4 groups together the results for the visual assessment: each column is associated with a different value of N ($N = 4, 6$ and 12 from left to right), the top row shows the empirical unit-variance cdfs together with the mesh-styled theoretical cdf of the zero-mean unit-variance joint bivariate Gaussian distribution, and the bottom row presents the fitting deviation magnitude[12] for each grid position on the plane. As expected, increasing N has a strong influence on the resulting 2D cdfs, whose resemblance with the theoretical normal cdf becomes remarkable for $N > 6$ (see $N = 12$ in 2.4c). This is confirmed by the evolution of the error magnitude, which becomes lower and more uniformly distributed with N. In numbers, the normal fitting accuracy is

[12]We define as $error^{ij} = |cdf^{ij}_{theory} - cdf^{ij}_{empirical}|$, where $i, j \in \mathbb{N}$ and correspond to grid positions on the plane.

expressed as the percentage of error magnitude reduction with respect to uniform distribution (i.e. 0% means perfect uniform distribution and 100% means perfect circularly-symmetric Gaussian distribution). The results are collected in Table 2.2, all in all showing that despite the dependencies between the quadrature components, the output distribution in SCM-PSM approaches capacity-achieving Gaussian as N increases (monotonically for $2 \leq N \leq 12$).

Table 2.2: Goodness of fit of PSM output distributions with respect to ideal Gaussian for $\{N \in \mathbb{N}_{even}; 2 \leq N \leq 12\}$. The values referenced to the ideal uniform distribution and presented in percentage fashion (0% means perfect uniform distribution and 100% means perfect Gaussian distribution).

N	Accuracy [%] (uniform ref.)
2	0
4	2.3
6	50.8
8	71.8
10	86.2
12	90.0

Shaping gain : Although the asymptotic convergence to normal-like distribution surely increases the achievable capacity under the considered assumptions, practical constraints prevent N from reaching arbitrarily large orders. Consequently, the goodness of fit is likewise bounded, diminishing the achievable gains now substantially lower than the *ultimate limit*. The study of the potential benefits of constellation shaping for technologically feasible modulation levels is thence an important step in system optimization and design. In the following, we compare the capacity of SCM-PSM against QAM in terms of receiver sensitivity and SE for $\{N \in \mathbb{N}; 4 \leq N \leq 9\}$, which encompasses the lowest modulation orders with positive Gaussian fitting (see Table 2.2). It is important to clarify that, in this study, the contribution to the channel noise comes from both linear and non-linear interference effects, which are treated as a single additive circularly-symmetric Gaussian random source. This consideration facilitates the inclusion of the power dependent distortions in the SNR parameter through an overall equivalent noise variance; and foremost, it

makes possible to extrapolate the mutual information results in the linear memoryless AWGN channel to the discussed optical transmission links.

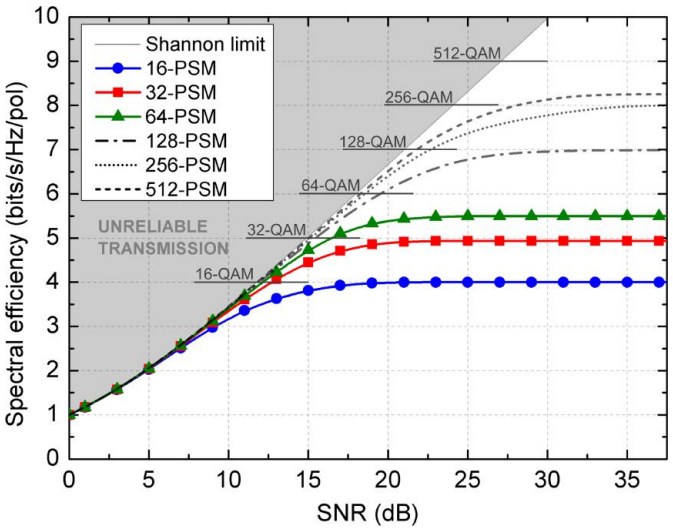


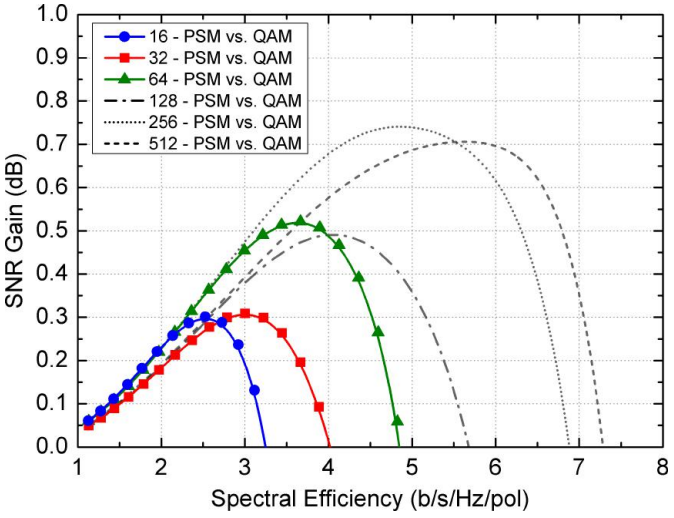
Figure 2.5: Spectral efficiency per polarization versus signal-to-noise ratio per symbol of PSM constellations for $\{N \in \mathbb{N}; 4 \leq N \leq 9\}$.

Figure 2.5 shows the constrained capacity of SCM-PSM $\forall N$ against SNR for memoryless AWGN channel, the maximum SE of QAM for various orders is included for reference purposes. As explained in Section 2.2.1.2 and clearly visible in the highest-SNR regions in Figure 2.5, non-bijective configurations (e.g. $N = 6$, see Figure 2.2b) cannot deliver the same maximum mutual information than same-order QAMs (always bijective). Note that these values of maximum absolute SE coincide with the theoretical information rates previously calculated and collected in Table 2.1. Based on the information provided by the capacity curves, shaping gains can be readily extrapolated by comparing the results for diverse modulation schemes. Accordingly, we first present the gain in SNR for different values of SE along with the associated shaping inaccuracies, providing thereby useful data for calculating the power budget increase given a target bitrate and modulation bandwidth. Secondly, the shaping gain is expressed in terms of SE for specific values of SNR, which could be employed for improving the channel code through the indicated extra redundancy.

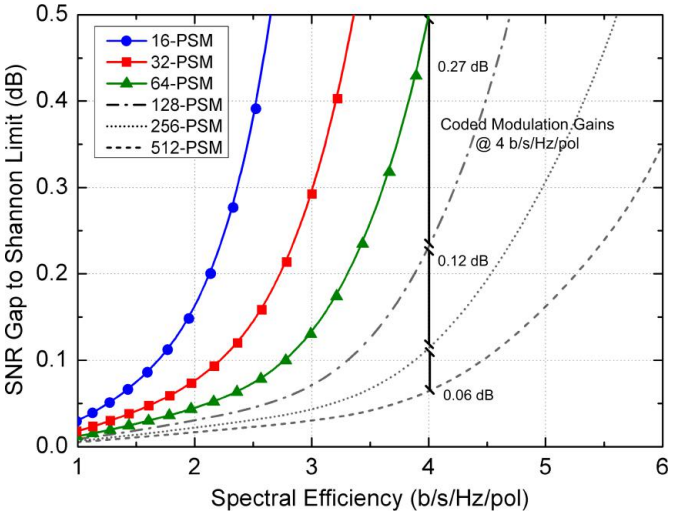
In Figure 2.6a, the amount of extra SNR degradation that SCM-PSM can endure over QAM for a range of SE values is shown $\forall N$. The information on this graph may be employed for calculating the transmission

reach extension of shaped signals for a pre-defined effective throughput. Because the shaped power distributions are designed for concrete channel responses, SCM-PSM consistently outperforms QAM in low-to-medium SE regions (peaking at ≈60% of N), where the performance is principally dominated by the channel noise. Beyond this threshold the gain decreases rapidly, noticeably faster than other geometric shaping techniques (see [71]) for certain N. This is partly caused by the ambiguities and the reduced Euclidean distance between symbols (e.g. Figure 2.2c) resulting from the irregularly spaced clusters in SCM-PSM constellations, which dictate the performance at high SNR. Notice that owed to the cross-like distribution of QAM for $N \in \{5, 7, 9\}$ (see Appendix A), the energy efficiency is slightly increased hence the comparable shaping gains between $N \in \{5, 7, 9\}$ and the immediate inferior even orders.

Complementing Figure 2.6a, the remaining SNR gap to channel capacity for SCM-PSM is depicted in Figure 2.6b. This graph quantifies the inefficiency (in SNR) of the output constellations as a result of their discrete nature and their imprecise Gaussianity (see Figure 2.4). Inferred from the non-crossing curves progressively approaching capacity with N, the first conclusion is that the gap to capacity closes $\forall$SE as N increases. In addition, it is noteworthy that despite the shaping gain increases and even peaks at specific SE values $\forall N$, the gap to capacity of SCM-PSM increases monotonically with SE. It follows that increasing the shaping gain does not mean closing the distance to channel capacity and, therefore, running the system at SEs where maximum shaping gains are attained may not be the most power efficient option. As an example, consider 64-ary PSM encoded with an ideal FEC with ≈35% total overhead. In that case, the SE (≈3.9 bit/s/Hz/pol) achieves maximum shaping (≈0.5 dB, see Figure 2.6a) though still missing ≈0.4 dB to Shannon limit (see Figure 2.6b). Given this configuration, setting the redundancy overhead to 50% (3 bit/s/Hz/pol equivalent SE) would reduce the shaping gain by negligible 0.06 dB while decreasing the gap to capacity by 0.3 dB extra (less than 0.15 dB remaining). These SNR savings add to the ≈3 dB obtained by accepting 1 bit of extra redundancy. Nevertheless, it is not often that we are willing to sacrifice effective throughput, so the next observation shall be regarded. At 4 bit/s/Hz/pol of SE, the shaping gains for 64-ary and 128-ary PSM converge to the same magnitude (≈0.49 dB), nonetheless, the gap to capacity is ≈0.27 dB lower for the second one (see superimposed intervals in Figure 2.6b) at the expense of ≈10% extra overhead. The process of combining higher-order constellation with stronger coding schemes, so as



(a) Shaping gain



(b) Gap to Shannon limit

Figure 2.6: Receiver sensitivity gain and remaining gap to Shannon limit of PSM constellations for $\{N \in \mathbb{N}; 4 \leq N \leq 9\}$.

to gaining extra sensitivity and at the same time maintaining the effective bitrate is called *coded modulation*.

An alternative interpretation of the benefits of constellation shaping is that, for the same SNR, higher information rate is possible as compared

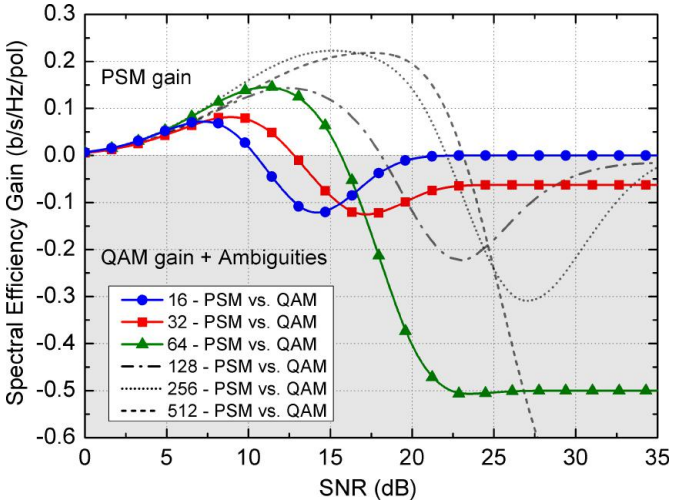


Figure 2.7: Shaping gain in SE for $\{N \in \mathbb{N}; 4 \leq N \leq 9\}$.

to non-shaped signals. A case where this information may be of pragmatic interest is when the transmission reach and link properties are given (approximately fixed SNR) and the error rate requisites to guarantee successful error correction are not met. The gain values of SCM-PSM versus uniform QAM $\forall N$ are presented in Figure 2.7 up to 35 dB of SNR. The top white-background part correspond to the area where PSM outperforms QAM, and *vice versa* for the bottom grey-background part. Previously hinted in the explanation of Figure 2.6a, the achievable performance, and then the shaping gain, is dictated by the ambiguities and the constellation grid irregularities for high values of SNR. This is apparent in Figure 2.7, where the falling-slope trajectories change with N exactly as the number of ambiguities and the constellation diagrams do. Note that the SE for those ambiguous modulation orders stabilizes below zero (e.g. $N = 5$ and 6), and it converges to zero gain in bijective cases (e.g. $N = 4$). Further information can be extracted from the existence of valleys and the depth of their minima after crossing the zero-gain point. If the gain curve describes fall-and-rise profile, it indicates that once the inter-symbol Euclidean distance start dominating the information rate, the regular grid of QAM is favorable in the corresponding SNR interval. In this line, the depth of the minima relative to the convergence gain tells about the severity of the grid-driven penalties. For example, both 16-ary and 32-ary PSM exhibit clear valleys in their profile, thus their grids entail error rate degradation

as compared to same-order QAM; however, the average Euclidean distance is larger for 32-PSM, given that the minima is the same for both orders, but not the convergence gain (0 and -0.06 bit/s/Hz/pol for 16-ary and 32-ary respectively). Especially attractive is the 64-PSM configuration, whose gain profile indicates that the shaping process does not penalize the Euclidean distance (as explained in Section 2.2.1.1) but only the information rate (see Table 2.1).

All these scattered examples demonstrate that shaping gain and power efficiency are dependent but not trivially related, requiring a meticulous optimization of the modulation parameters in which the coding system (indeed never ideal) plays a fundamental role[13]. Particularly in SCM-PSM, the coding approach is not only fundamental for maximizing the power efficiency but also to resolve the ambiguities caused by the mapping/shaping architecture that, in short, prevent the successful transmission of uncoded data irrespective of the SNR.

2.2.2 Detection and de/coding

As previously discussed, non-bijective symbol arrangements may occur for certain constellation orders during the mapping procedure in SCM-PSM. Even though we have shown that the information rate does not decrease by more than 9% (see Table 2.1), successful signal demodulation is subjected to the complete mitigation of the ambiguous inter-layer interferences. This task becomes further challenging owed to the minute Euclidean distances between some non-overlapping symbols, aggravating the uncertainty under the low-to-medium SNR conditions in which shaped constellations are intended to operate (this holds for any power-constrained non-equally spaced constellation grid). Consequently, highly inaccurate demapping decisions are initially made, strongly suggesting iterative belief propagation between demapper and decoder. The fundamental idea is that both blocks repeatedly exchange Bayesian information to mutually improve their extrinsic information, thus progressively refining symbol detection, ambiguity resolution and error correction. Adjusting the parameters of such iterative SISO receiver to efficiently use up the available shaping gain, demands meticulous study of the extrinsic information flow and its dependence with SNR, inter-symbol distance, symbol labeling and encoding scheme. In spite of its conceptual simplicity, the latter is a convoluted optimization process

[13]Implementation penalties, processing latency and demodulation robustness are other aspects that determine the actual system design.

with big impact on the processing complexity and the error correction capabilities.

In this concern, we next elaborate on SCM-PSM from the digital receiver perspective, subdividing the analysis into (i) coded modulation structure (ii) demapping and (iii) channel code. The essential properties of SCM-PSM regarding the design of the iterative SISO receiver and the associated channel code are discussed, as well as general considerations of practical interest.

2.2.2.1 Coded modulation and iterative SISO receiver

Two main reasons motivate the use of bit-oriented coded modulation together with PSM. First, PSM works autonomously when the superimposed layers are statistically independent and present uniform data distribution, thus generating asymptotically accurate 2D Gaussian-like constellations independently of the FEC and encoding strategy. Second, when linear mapping[14] is used, symbol labels are typically built by concatenating the binary values of the N conforming branches (one bit per layer). Even if different bit-significance conventions are employed, the mapping function remains systematic according to PSM equations (see 2.3), compromising the labeling freedom and increasing the optimization complexity. That is, the adoption of a bit-oriented coded modulation framework is encouraged for facilitating seamless integration of the PSM shaping module, and also the joint de/coding and de/mapping optimization in spite of the rigid labeling.

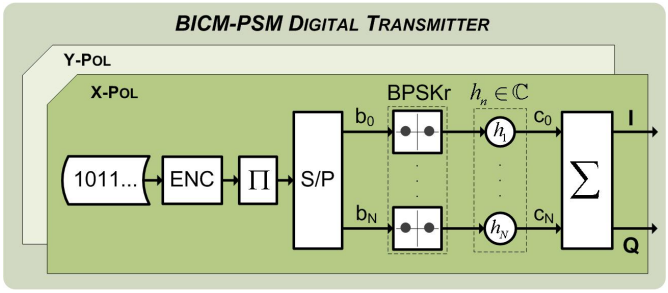
In this regard, bit-interleaved coded modulation BICM has been widely proved as enabling bit-oriented approach whose flexibility allows for the coherent application of powerful families of binary codes on virtually any modulation format [76]. The respective extension of BICM to support iterative decoding, BICM-ID[15] [77], was employed in this investigation. It

[14] We refer to linear mapping as the process of generating the symbols based on the linear superposition of weighted antipodal values. In contrast, non-linear mapping is the process of assigning bits to a pre-defined constellation (i.e. 16-ary QAM) according to certain rules that do not hold a linear mathematical relation with the input bits.

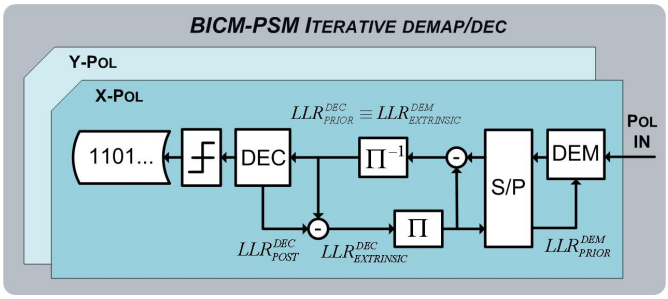
[15] A posteriori probabilities (APPs) in the form of log-likelihood ratios (LLRs) per coded bit are repeatedly circulated among the demapper and decoder. These factors are calculated as:

$$LLR_{POST}(c_m) = \log \frac{p(y|c_m = 0)}{p(y|c_m = 1)} \tag{2.8}$$

where c_m is the mth coded bit in the block being processed, and $p(y|c_m)$ is the conditional



(a) Transmitter



(b) Receiver

Figure 2.8: Digital transmitter and SISO iterative receiver of BICM-PSM.

should be noted that SCM with single-layer encoding strategy and bit-wise interleaving is identical to BICM with superposition mapping. For reading ease, and without loss of generality, we refer to SCM with single-layer encoding strategy, bit-wise interleaving and PSM as BICM-PSM. Figure 2.8 shows the the schematic of the digital transmitter and iterative receiver in BICM-PSM.

probability of y given c_m. The extrinsic (new) information is extracted from the generated APPs by removing the fraction of old a priori content (LLR_{PRIOR}):

$$LLR_{EXTRINSIC} = LLR_{POST} - LLR_{PRIOR} \tag{2.9}$$

Then, the $LLR_{EXTRINSIC}$ are passed to the next block (demapper towards decoder or *vice versa*) as new a priori information. This continuous dialog progressively increases the reliability of the transferred APPs and, in turn, the decoding accuracy and bit error rate (BER) performance. After a predefined number of iterations, the LLRs of the information bits are fed into a slicer that outputs the estimated binary sequence.

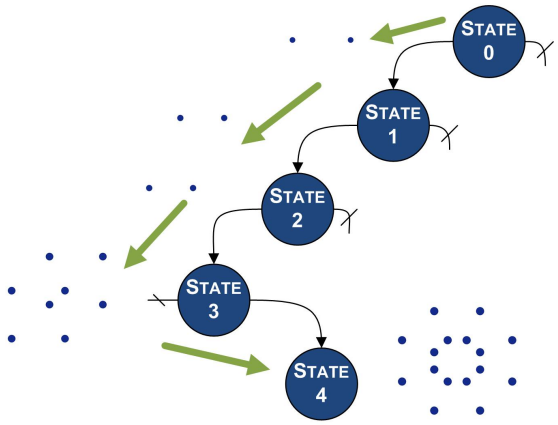


Figure 2.9: Path evolution on the tree-diagram model of PSM. The constellation diagrams for every state > 0 are included.

2.2.2.2 Demapping

In relation to the algorithms for SISO demapping, MAP methods are extensively applied as they give optimal estimation of the states, or outputs, of a Markov process observed in AWGN [78]. Bridging to our case of study, after signal propagation in the target scenarios, optimum symbol estimation is achieved if the mapping procedure behaves as a Markov process. This is the case of superposition modulation, whose operation can be modeled as a tree diagram [9] in which two branches emerge from every state, both accounting for each of the possible outputs of the weighted BPSK symbol (see Figure 2.9). Despite attaining optimal layer separation and demonstrated intra-channel nonlinearity mitigation [79], general MAP symbol estimation suffers from scalability limitations due to the quadratic dependence of the computational complexity with N. For this reason, various approaches have been presented in the literature so as to alleviate this circumstance. For instance, seizing upon the fact that superposition modulation modeled as a Markov process, the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [80] can be used instead of the general MAP estimator to strongly reduce the computational load and processing latency without performance degradation.

The estimation accuracy of the average transition probability of the channel is another significant factor to muse over when employing SISO demapping. As argued in previous sections, the combined imprint of both linear and non-linear impairments after propagation in the scenarios re-

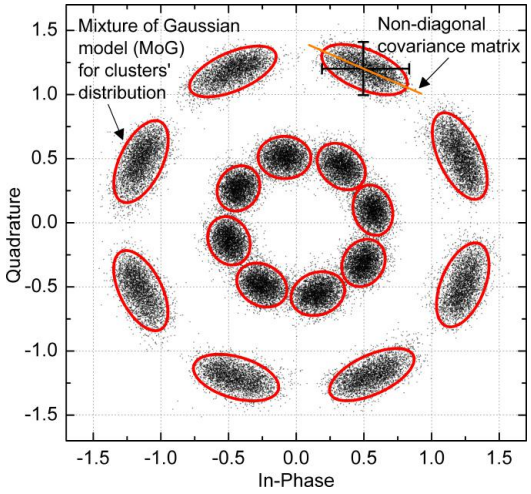


Figure 2.10: Recovered 16-ary PSM constellation after 240-km uncompensated SSMF transmission at 6 dBm launched power. The per-cluster covariances estimated by the expectation maximization algorithm are shown at 99% confidence.

garded in this chapter is typically simplified to additive and circularly-symmetric Gaussian noise. Whereas this approximation has been widely adopted, its validity has been likewise proved positively gradual with the transmission reach (lumped amplification and WDM technology are invariants), originating an apparent noise correlation for insufficient propagation distances. Albeit this fact would have an impact in symbol detection for any constellation diagram, the extreme proximity of the symbols in the lower power areas in PSM output constellations makes the accurate calculation of per-cluster centroids and covariances especially useful. In this investigation, we make use of a two-step iterative procedure called expectation maximization (EM) [81] for optimum estimation in a maximum likelihood sense (see Figure 2.10). The calculations are simplified by assuming a mixture of Gaussian (MoG) statistical model for each cluster, and restricting the MoG to a 2D space (2x2 covariance matrix). See [82] for details on EM algorithm applied to this task.

2.2.2.3 Channel coding

Indispensable for providing the required coding gain (virtually increasing the Euclidean distance among symbols) and diversity gain (for ambigu-

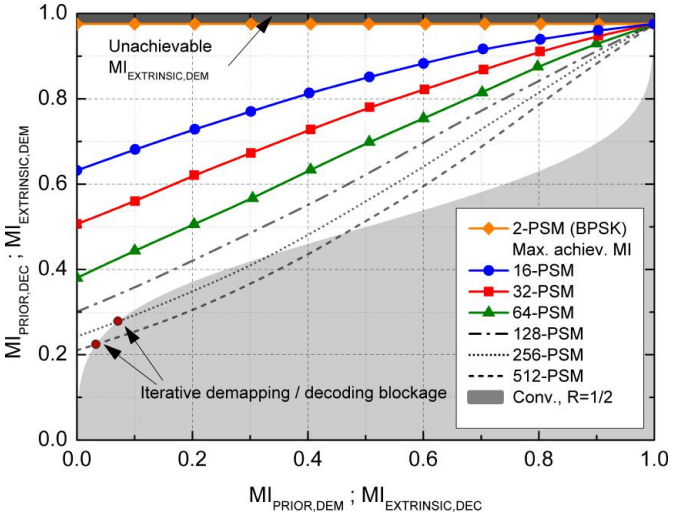


Figure 2.11: EXIT profiles of PSM soft demapper for $\{N \in \mathbb{N}; 4 \leq N \leq 9\}$ at 8 dB SNRb, and BCJR decoder for 0.5 channel code rate convolutional code.

ous multi-layer interference resolution), the election/design of the adequate code for ID receivers is inseparably linked to the modulation properties, the SNR and the requirements of the application (e.g. processing latency or bit error tolerance). It follows that claiming the ubiquitous superiority of any code is inexact. Joint optimization of the transmitter and receiver parameters for pre-defined SNR values is required to minimize inefficiencies in the SISO demodulation and thus assess appropriateness of the proposed coded modulation architecture. This rigorous process starts observing the evolution of the extrinsic information transfer flow between demapper and decoder, intuitively carried out with the help of extrinsic information transfer (EXIT) charts (see [83] for detailed information).

In Figure 2.11, the EXIT profile of the PSM BCJR demapper for $\{N \in \mathbb{N}; 4 \leq N \leq 9\}$ is depicted at 8 dB of SNRb together with the one of BCJR decoder for 50% overhead (0.5 channel code rate) convolutional code. Due to the high symbol density in the lowest power area of Gaussian-like geometrically shaped constellations, the MI from the demapper in the absence of a priori information ($MI_{PRIOR;DEM}$) is considerably weak as compared to uniform distributions . Particularly in PSM, the symbols can be arbitrarily close or even overlap (ambiguous mapping) resulting in EXIT functions that become convex with the modulation order (clearly visible

for $MI_{PRIOR;DEM} < 0.4$). Already pointed out previously, this situation forces the decoder to generate strong extrinsic information based on the hardly reliable initial a priori MI from the demapper in order to prevent the blockage of MI improvement at an early stage (examples of blockage are illustrated as red points for $N = 9$ and 8). This represents a major restriction imposed, to larger or smaller extent, by all power-constrained geometrically shaped constellations (see [71]) and makes of zero-iteration receivers an inappropriate choice.

Given that PSM symbols are built from the successive accumulation of phase-shifted antipodal values, the MI profile for all the modulation orders converges to that of BPSK when decoding reliability tends to 100%. Notice that an alternative interpretation of the superposition of branches is that $N - 1$ addends act as deterministic interference channels for the remaining BPSK-modulated one. When this interference noise is compensated, equivalent performance to single-branch transmission is accomplished independently of N and the level of SNR. This has an important implication: since the maximum MI out of the demapper will always end at a high value, the profile of the decoder can similarly be designed to end at a very high point. This means that, unlike in high-order QAM for instance, strong coding gains[16] are not strictly necessary in order to achieve complete error correction given sufficient number of iterations. In addition, because maximum output mutual information from the demapper is steady $\forall N$ and fixed SNR, the transfer function will progressively steepen as the N-dependent Euclidean distances reduce. This emphasizes the need for iterative receivers under the assumption of well-fitted demapper-decoder EXIT functions, which is the idea behind efficient code design.

Words on selected code : Even though each modulation order requires specific code design, for the proof of concept a 0.5 channel code rate convolutional encoder with $(5, 7)_8$ polynomial generator was implemented for the forthcoming experimental demonstration. The BCJR simplification of the MAP decoder is implemented for optimum performance [78]. This scheme enables complete layer separation within 10 iterations for the three experimentally studied cases, and sets the number of effective bit/symbol very close to the maximum shaping gain regions (see Figure 2.6a).

[16]The MI that the decoder is able to deliver to the demapper for strongly reliable prior extrinsic information

2.3 Experimental demonstration

Up to this section, BICM-PSM has been revisited from the point of view of digital communications and information theory. It has been concluded that, in channels where the noise statistics within the channel of interest can be modeled as memoryless additive Gaussian, BICM-PSM can effectively approach capacity for sufficiently high modulation levels when the coded modulation scheme is designed/adjusted accordingly. These promising gains, however, assume that signal generation and recovery are, not just possible, but ideal. This is truly harsh conjecture in coherent optical communications, further emphasized by the potential difficulties that irregular constellation grids pose on blind phase-noise compensation and channel equalization. The ensuing implementation penalties jeopardize the sensitivity gain provided by the shaping process in different manners. In the transmitter, insufficient quantization resolution (when DACs are used) or non-linear electro-optical conversion make the output power distribution divert from the optimum one. During propagation, varying the conditions that support circularly-symmetric additive Gaussian noise approximation will reduce the shaping effectiveness. In the receiver, non-linear effects during field recovery and low digitization resolution will likewise distort the optimum shape and thus the associated gain. Finally, unsuitable DSP algorithms may produce noise undercompensation or enhancement thus diminishing the effective SNR coming into the SISO receiver, formerly conceived for optimum operation in a different working point. Because of all the above-mentioned aspects, studying the experimental feasibility of BICM-PSM is indispensable for correct end-to-end power budget calculation. This is the context of the following section, where we have focused on the evaluation of the implementation penalties with respect to ideal digital transmission when standard DSP algorithms are used for signal recovery and impairment compensation. This study represents the just a first steps towards complete characterization.

Dual-polarization single-carrier 16-ary, 32-ary and 64-ary PSM are experimentally generated, transmitted and recovered. Numerical simulations of digital transmission over purely AWGN channel for coded and uncoded configuration are included for benchmarking optical B2B measurements. Theoretical, simulated and experimental effective SEs are compared and discussed for all the cases of study. Regarding transmission, we benefit from EM-based cluster parametrization to achieve a slight sensitivity improvement after non-linear propagation over 240-km uncompensated SSMF.

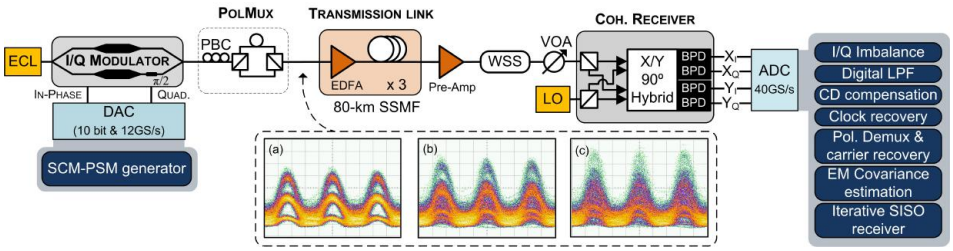


Figure 2.12: Schematic of the experimental setup for dual-polarization single-carrier BICM-PSM transmission. Insets: optical eye diagrams for $N = 4$ (a), $N = 5$ (b) and $N = 6$ (c).

2.3.1 Testbed

The schematic of the experimental setup is shown in Figure 2.12 along with root-raised-cosine (RRC) pulse-shaped DP eyediagrams for $N \in \{4, 5, 6\}$, which were captured after ≈10-s accumulation time. As a hint of the shaping effect, it is distinguishable that the average point density concentrates around the lowest power range as N increases. In the transmitter side, a 1550-nm external-cavity laser (ECL) with ≈100 kHz linewidth is used as the light source. The output of the laser is modulated by a dual-drive optical I/Q modulator, which is differentially driven (≈2 Vpp) directly from the outputs of a 10-bit resolution arbitrary waveform generator (AWGen) employed for electrical signal generation. The sampling rate and the baudrate are fixed over the entire experiment at 12 GS/s and 6 GBd (2 S/symbol) respectively. A pseudo-random bit sequence (PRBS) of length $2^{15} - 1$ were digitally processed to generate the waveforms according to Figure 2.8a. The digital signal processing consists of: encoding, interleaving, mapping, upsampling and pulse shaping. First, 150000 bits are encoded block-wise in a single-level strategy with a 0.5 rate convolutional code and $(5, 7)_8$ polynomial generator; afterwards, half-random interleaving is realized (lengths vary within 10000-15000 bits depending on the modulation level) to break the sequential correlation and increase diversity order. The sequence is rendered parallel, offering further scrambling and preparing the data for PSM (see Subsection 2.2.1.1 for details about 1 S/symbol signal generation). In the last digital block, 50 quadrature phase-shift keying (QPSK) pilot symbols are inserted every 1000 data symbols (5% overhead) for estimating the channel response and compensating the frequency offset from coherent intradyning; finally, the aggregated waveform is 2x oversampled and pulse-shaped with 0.8-roll-off factor (β) RRC.

After the optical I/Q modulator, dual-polarization multiplexing is emulated; the transmitter output is amplified and then launched into an uncompensated fiber link for transmission. The link is made up of 3 spans of 80-km SSMF and Erbium-doped fiber amplifiers (EDFAs) for compensating the attenuation after each span.

In the receiver, standard coherent detection is carried out. The incoming optical data signal is coherently mixed with the local oscillator (LO) (ECL, ≈100 kHz linewidth) and, subsequently, the outputs from the four balanced photodiodes are digitized at 40-GS/s and stored for offline processing in a digital storage oscilloscope (DSO) with 13-GHz analogue bandwidth. The receiver DSP comprises I/Q imbalance compensation, CD compensation, low-pass filtering, clock recovery, joint polarization demultiplexing and equalization, carrier frequency and phase recovery and SISO BICM receiver with ID.

DSP workflow and punctualizations : After standard front-end correction and frame synchronization, the 10-tap filters of the adaptive equalizer are first estimated with constant-modulus algorithm (CMA) [84, 85] operating on the QPSK pilot sequence. After pre-convergence, PSM-optimized radius-directed equalizer (RDE) [85,86] (sometimes referred to as multi-modulus algorithm (MMA)) is used for fine tap adaptation[17]. Note that the circular symmetry that PSM constellations exhibit results in fewer discrete constant-magnitude levels than in the analogous square-QAM $\forall N$, hence showing faster convergence rate for the same number of taps and SNR conditions. Table 2.3 presents the number of discrete power levels for both QAM and PSM up to $N = 12$, as well as the multiplicative scaling ratio associated with PSM for each N that makes the minimum distance between magnitude levels be the same as in QAM[18].

Next, carrier recovery is realized with second-order digital phase-locked loop (DPLL) [87], where minimum Euclidean distance metric is used. The proportional loop gain is minutely adjusted for each modulation order and SNR level. The information bits are recovered after 10 SISO iterations of error correction plus demapping optimization. For the iterative receiver to

[17]Despite knowing the QPSK sequence beforehand enables 2-rank complex channel matrix calculation and inversion, blind CMA was chosen instead for reasons related to code integrability. After running RDE, the equalizer filters' taps showed less than 5% average discrepancy between both methods in MMSE sense. The implications on the computational complexity are disregarded.

[18]Larger than 1 means that the minimum distance between magnitude levels/rings in QAM is smaller than in PSM and *vice versa*.

Table 2.3: Number of constant magnitude levels in QAM and PSM constellations for $\{N \in \mathbb{N}; 4 \leq N \leq 12\}$. The rightmost column shows the scaling factor to be applied on QAM in order to equalize the obtain similar minimum distance between constant magnitude levels than PSM constellation for a given N. A value larger than 1 means that the minimum distance between constant magnitude levels is smaller in QAM by as much as the respective value indicates.

	# of Mag. levels		
N	QAM	PSM	Min. Inter-ring distance ratio ($\frac{\text{QAM}}{\text{PSM}}$)
4	3	2	2.24
5	6	4	2.89
6	9	5	4.02
7	20	9	3.75
8	32	12	2.86
9	56	17	2.52
10	109	34	2.89
11	235	63	1.13
12	398	68	4.77

work optimally, the demapper is fed with maximum likelihood covariance estimations calculated through the EM algorithm (see Subsection 2.2.2.2). This exercise boosts the performance of the ID receiver in situations where the average pdf varies between clusters like, for instance, when random phase noise is not totally compensated by the DPLL. In our case, phase noise is intentionally introduced via self-phase modulation (SPM) for analysis purposes. Whenever the average transition pdf at the input to the SISO receiver is well approximated by circularly symmetric additive noise, per-cluster pdf estimation is meaningless and shall be sidestepped.

2.3.2 Results

The discussion on the results is subdivided into (i) optical B2B, (ii) spectral efficiency, and (iii) transmission.

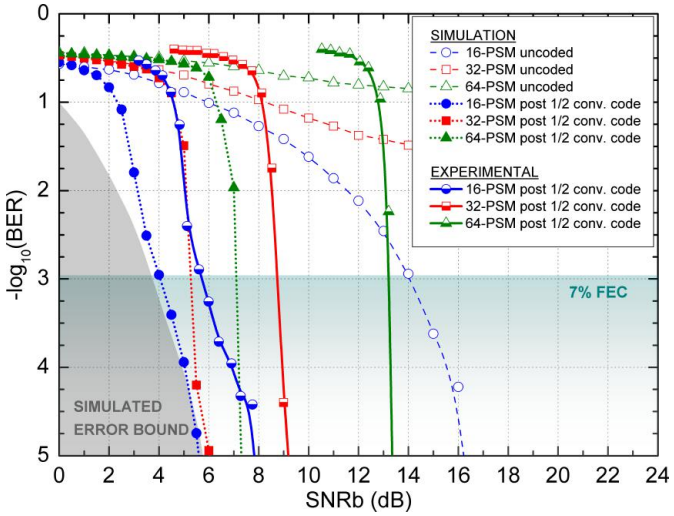


Figure 2.13: BER as a function of SNR per information bit for optical back-to-back measurements and simulated digital transmission over AWGN channel. Coded and uncoded configurations are considered.

2.3.2.1 Optical back-to-back

Figure 2.13 shows BER as a function of SNRb for measured optical B2B after soft-decision decoding for 16-ary, 32-ary, and 64-ary PSM. Numerically calculated BER curves for coded and uncoded PSM after AWGN-only channel transmission are included as reference. All the BER measurements result from averaging three samples of 150000 bits each. Accounting for the possibility of using an additional outer decoder in a concatenated coding/decoding structure [28], the pre-FEC BER threshold ($1.1 \cdot 10^{-3}$) of a 0.9375 code rate Reed-Solomon code [29] is traced for analysis purposes. The grayed area represents the numerically calculated BER boundary for BICM-PSM. This error bound corresponds with the performance of 2-ary PSM (BPSK), and it is independent of the number of iterations and the modulation order (see Subsection 2.2.2.3).

As previously explained, the occasional information rate reduction in PSM's constellations requires using a proportional percentage of redundancy that enables sequence corroboration; otherwise, data recovery is impossible independently of the level of SNR. This is apparent in the uncoded results for 32-ary and 64-ary PSM, whose ambiguities cause BER floors around $3.2 \cdot 10^{-2}$ and $1.3 \cdot 10^{-1}$ respectively. Demodulation is pos-

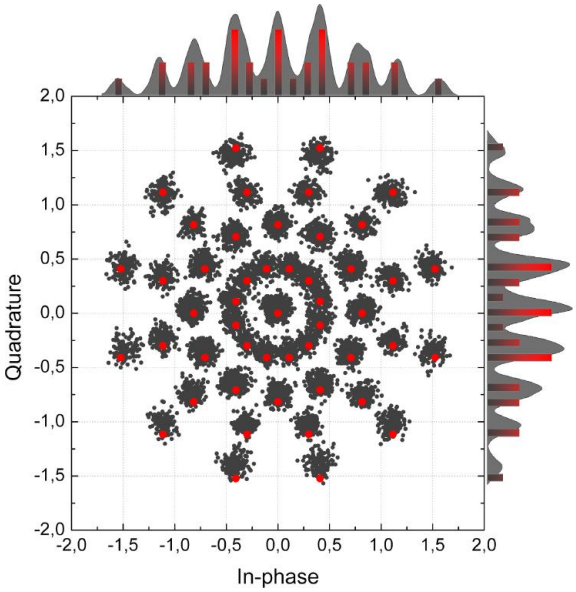


Figure 2.14: Average-power normalized experimental constellation diagram of 64-ary PSM for ≈30 dB OSNR (gray). The ideal symbol distribution is superimposed (red). The histograms for each quadrature of both data sets are included at the corresponding sides of the measurement box.

sible with SISO iterative decoding , showing strong coding gains (≈10 dB for 16-ary PSM at the outer decoder pre-FEC BER threshold) with turbo cliffs starting at 2 dB for $N = 4$, 4 dB for $N = 5$, and 6 dB for $N = 6$ for the numerically simulated data. Experimental data exhibits similar behavior, showing the first successful demodulation and decoding of PSM signals in coherent optical communications. This results confirm that ambiguities and transmission errors can be simultaneously resolved employing BICM structure with ID, as long as the channel code rate is higher or equal than the constellation's information rate (see Table 2.1) and the EXIT profiles of both demapper and decoder do not cross at the desired SNR level [21].

Even though coded modulation makes demodulation possible for increasing SNR, it is visible that simulated coded 16-ary and 32-ary PSM experience error correction limitation starting at ≈3.5 dB and ≈5.5 dB SNRb respectively. The reasoning is intuitive from the analysis of the EXIT chart (see Figure 2.11). As SNRb decreases beyond certain threshold, PSM's demapper EXIT profile starts moving downwards to cross with the decoder's (static with SNR) at a point where $MI_{EXTRINSICDEM} < 1$,

hence preventing complete error correction irrespective of the number of iterations (see the convergence towards the system's error-rate boundary). Note that, in contrast to the demapper EXIT functions of the three considered cases, the convolutional decoder evolves from concave to convex as the input MI increases. This causes that the iterative exchange of extrinsic information gets blocked at an early stage for very low SNRb, thus producing worse BER than the system's minimum error rate (clearly visible in 16-ary PSM for 0 dB $\leq$ SNRb $\leq$ 3.5 dB). These observations illustrate the importance of joint demapper-decoder optimization in coded modulation schemes for maximum computational efficiency and error correction capabilities. Regarding experimental curves, owed to implementation penalties (1.7 dB, 3.3 dB and 6 dB for 16-, 32- and 64-PSM at the outer decoder pre-FEC BER threshold) requiring higher SNRb to obtain similar BER, such error floor is not visible anymore for 32-ary PSM, whose turbo cliff starts at 7.5 dB SNRb. However, the tendency of the boundary is strictly followed in the 16-ary PSM case, where the implementation penalty is not so severe. Realize that implementation penalties change the demapper EXIT profile for every N. In all cases, the major contribution to implementation penalties comes from the DPLL not totally compensating for the lasers phase noise. Notice in Figure 2.14 the clear separation of magnitude levels (see Table 2.3 and comments in Subsection 2.3.1), and that hardly any distortion is appreciated in the symbol locations with the exception of a slight compression in the quadrature component. Despite the DPLL loop parameters are optimally adjusted for every measurement point and the use of EM-based per-cluster pdf estimation, the low baudrate together with the stringent SNR conditions in the maximum shaping gain ranges pose clear sensitivity degradation.

2.3.2.2 Spectral efficiency

Figure 2.15 shows the conversion to effective SE per polarization versus SNRs for the BER results at the outer FEC BER threshold ($1.1 \cdot 10^{-3}$) in Figure 2.13. The BER to SE calculations are conducted under the assumption of brickwall power spectral density (PSD) for all the cases (see [88] for details on the conversion method). Maximum SE-wise shaping gain regions for $N \in \{4, 5, 6\}$ are shaded within the corresponding SNRs intervals (see Figure 2.7), and effective SE curves of the simulated BER results after inner decoding for $N \in \{4, 5, 6\}$ (discontinuous lines) serve as benchmark. Full-filled symbols represent the effective SE of each of the simulated cases under study at above-mentioned pre-FEC BER threshold. The SE values

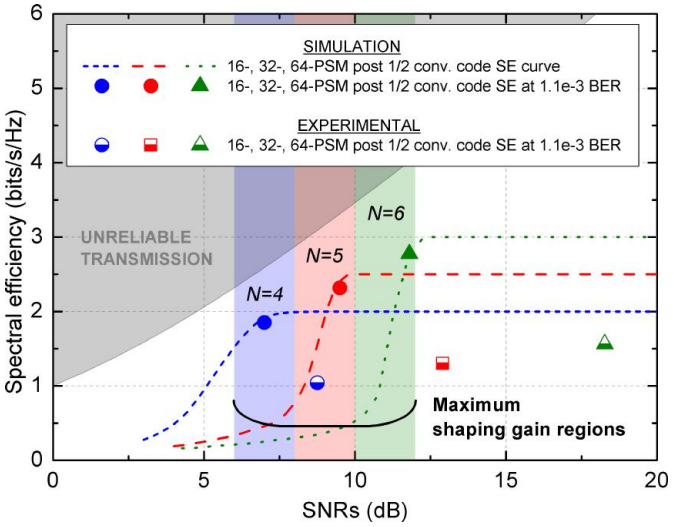


Figure 2.15: Effective SE as a function of SNRs for optical back-to-back measurements and simulated digital transmission over AWGN channel. Color-matched shades indicate maximum shaping gain SNRs intervals for each case.

account for both the 0.5 and 0.9375 channel code rates of the inner and outer decoder respectively ($0.5 \cdot 0.9375 = 0.46875$ equivalent channel code rate). As expected from the analysis in Subsection 2.2.1.3 (see Figure 2.6a), the receiver sensitivity for all the modulation orders fall within the ascribed areas of maximum shaping gain for ≈50% effective information rate. This means that the coded modulation configuration elaborated in Subsection 2.2.2 plus the suggested outer FEC decoder can perform at $< 10^{-15}$ BER with a SNRs tolerance that falls within SE-wise maximum shaping gain intervals $\forall N$. However, it should be noted by comparing the simulated results against those in Figure 2.6b that the achieved SE is clearly apart from the capacity limit for all the cases. Taking as example 64-ary PSM, we see in Figure 2.15 that gap to capacity is ≈4 dB, whereas the theoretical results in Figure 2.6b indicate that the achievable capacity at the aggregate information rate (≈2.81 bit/s/Hz/pol) is approximately 0.1 dB. This means that, even though the the concatenated code structure makes possible to tolerate the stringent SNR levels at which shaping gain is maximized, the efficiency of the 0.5 convolutional code at the target SNR from the EXIT perspective (understood as how well it fits the demmapers') is very low. This observation reemphasizes the importance of the joint opti-

mization of the coding/decoding strategy and the modulation format, and totally discards the combination of 0.5 code-rate convolutional codes with PSM.

For their part, half-filled symbols stand for the effective SE of the experimental results at the outer decoder pre-FEC BER threshold. Unlike simulated curves where the PSD was assumed rectangular ($\beta \to 0$), experimental values account for $\beta = 0.8$ pulse-shaping in addition to the 0.46875 overall channel code rate. Therefore, the effective SE per polarization is given by $N \cdot 0.46875 \cdot \frac{6\ [\text{GBd}]}{6\ [\text{GHz}] \cdot (1+\beta)}$. The results prove that implementation penalties can increase the necessary receiver sensitivity well beyond the areas of shaping gain. Considering that most penalties are caused by incomplete/incorrect phase noise compensation, we conclude that standard second-order DPLL is not an appropriate algorithm for phase noise compensation in the scenarios of interest when shaping gain is pursued; suggesting the use of more robust alternatives such as data-independent frequency pilot tones [46].

2.3.2.3 Transmission

Figure 2.16 presents the measured BER as a function launch signal power for 16-ary, 32-ary and 64-ary PSM after 240-km uncompensated SSMF transmission. Recovered constellations with EM-based covariances for each cluster are shown as insets. The analyzed input powers range between 0 dBm and 8 dBm in steps of 0.5 dB, and the BER values are averaged over three traces of 150000 bits each. Zero counted BER transmission is measured below 2.5 dBm, 6 dBm and 7.5 dBm for 64-ary, 32-ary and 16-ary PSM respectively. Because of the insufficient accumulated dispersion, SPM manifests as phase noise on the constellation diagrams. The resultant ellipticity and the noisy clusters and their centroids is correctly tracked by the EM algorithm hence improving the performance of the ID routine and making demodulation possible at considerably high launched power levels (up to 7.5 dBm for 16-ary PSM). Compared to the approach when the SNR is assumed circularly symmetric Gaussian with variance estimated from the worst-case cluster, the receiver sensitivity experiences an average improvement of 0.6 dB at the expense of $\approx (2^N - 1)$ times more complexity in the pdf calculation stage. The previous observation denotes the local imprecision of the assumption about the circular symmetry of the noise, and simply suggests considering non-Gaussian constellation shapes.

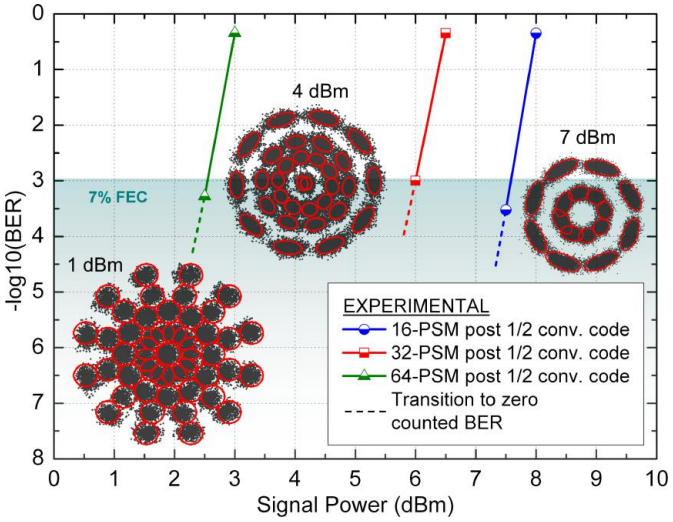


Figure 2.16: BER as a function of launched power after 240-km uncompensated SSMF transmission. Insets: Recovered constellations with the estimated covariances per cluster for $N = 4$ (7 dBm, right), $N = 5$ (4 dBm, middle) and $N = 6$ (1 dBm, left).

2.4 Summary

Current long-haul networks need higher throughput-times-distance product. The underlying limitation results from non-linear effects (random and deterministic unknown) opposing to the power-driven SE gains that technologies like AMF or grid-less WDM provide to these finite-bandwidth links. Urged by the need for more throughput and supported by the progress in digital processing capabilities, manifold DSP techniques aiming the compensation and/or mitigation of such fundamental nonlinearities are being intensively looked into. In this context, coded modulation featuring constellation shaping is deemed an attractive approach to complement existing compensation techniques for further sensitivity improvement.

Given our band-limited power-constrained channel, molding the power distribution of the transmitted signal can close the gap to capacity. Unfortunately, such optimistic goal is subjected to the perfect knowledge of channel's transition pdf, which is computationally intractable; and the use of infinitely large constellation size (continuous input), which is impractical from the implementation point of view. Consequently, it is accustomed to approximate the overall signal interference as additive power-dependent

2D Gaussian noise as well as the use of pragmatic modulation orders. All in all, the achievable capacity is constrained. Even if the receiver sensitivity is still improved, the shaping gains are diminished to a point where the complexity-to-gain ratio turns unfavorable; then bringing the interest towards constellation shaping methods that exploit the assumptions to simplify the process.

In this regard, SCM had been proposed as scalable and flexible coded modulation scheme that allows for considerable complexity reduction owed to its inherent linear bit-to-symbol mapping. In combination with PSM, SCM can (i) passively generate circularly-symmetric Gaussian-like constellations that asymptotically approach (with modulation level) memoryless AWGN channel capacity while maintaining high information rate; and (ii) simplify the general MAP demapper which, over and above, becomes optimum as symbol generation in SCM can be modeled by Markov process. These reasons motivate the consideration of SCM-PSM in the context of long-haul scenarios, as well as studying its information transfer capabilities, and the experimental assessment of its implementation feasibility in coherent systems.

This investigation

In this chapter, SCM-PSM is revisited as a promising coded modulation scheme with constellation shaping capabilities for increasing the achievable capacity of LH and ULH optical transmission systems at lower complexity as compared to equivalent techniques. In the first block, signal generation and detection are thoroughly covered. The analysis shows that PSM passively generates asymptotically accurate 2D Gaussian constellations with less than 8.4% information rate penalty (minimum 0%) for up to 12 bits per symbol. Resorting to the Gaussian noise model, PSM is thus able to increase the achievable capacity of the considered scenarios. These mapping ambiguities have been proved beneficial in PSM for certain modulation orders (i.e. 6 bit/symbol), where the average Euclidean distance among symbols is higher than in QAM when the average power is equalized. The asymptotic convergence rate towards Gaussian distribution as well as the fitting inaccuracies are theoretically characterized based on the calculated achievable shaping gains (both in terms of SE and SNR) and the absolute gap to Shannon limit. The ranges where PSM outperforms QAM are reckoned, in turn establishing the base for the initial decisions on the code strategy and the shaping-wise optimum transmission reach per modulation order.

As far as the coded modulation scheme is concerned, the use of bit-wise approaches employing MAP detection is well justified in combination with PSM. Coded signal generation and detection plus decoding is hence carried out with BICM (SCM with single-layer encoding) with SISO ID receiver. Based on the calculated maximum shaping gain ranges, 0.5 code rate convolutional code is employed for its simplicity and the optimal decoding performance when BCJR algorithm is employed. Likewise modeled as a Markov process, PSM operation enables the use of BCJR demapper instead of general MAP, on the whole simplifying the iterative receiver's processing load. With this framework, the criteria for efficient code design are then discussed with the help of the resulting EXIT chart. Among the most relevant conclusions we note that efficient decoding of PSM constellations demands iterative routines, and that PSM requires lower coding gains than QAM to achieve successful demodulation irrespective of the labeling.

In the second and last block, the experimental validation of the previously analyzed BICM-PSM scheme with SISO ID receiver is presented. Coherent detection is employed to recover single-carrier DP 6-GBd (gross-rate) 16-ary, 32-ary and 64-ary PSM constellations in B2B configuration and after uncompensated 240-km SSMF transmission with lumped amplification. The contemplated figures of merit include BER and effective SE. The SE results show that even though complete error correction and ambiguity resolution after $\leq$10 iterations is obtained, 0.5 code-rate convolutional codes are highly inefficient in combination with PSM despite the relatively similar EXIT profiles; penalizing the gap to capacity by more than 3 dB. Thorough matching of the EXIT functions of both decoder and demapper at the target SNR is mandatory for designing dedicated codes. The experimental measurements in B2B configuration hint the need to improve the robustness of the algorithm for phase noise compensation under low SNR conditions, which represented the main contribution to the implementation penalties. In this line, per-cluster covariance an centroid estimation using EM algorithm was proposed and experimentally demonstrated to reduce the impact of phase noise on the performance of systems employing soft-based receivers. By feeding the demapper with optimally estimated cluster parameters in maximum likelihood sense, the average receiver sensitivity was improved by $\approx$0.6 dB in highly non-linear transmission regime (SPM-dominated phase noise imprint).

Future work

The present investigation conveys the following items:

- Dedicated code design for PSM constellations : optimal code elaboration is an necessary step to truly exploit shaping gains and close the gap to constrained capacity down to the theoretical limits [5]. Comparisons with most popular modulation formats or alternative shaping models would then be possible. This is the way to make PSM worth consideration for practical implementation.

- Multi-layer coding strategy plus PSM: coding gains could be further improved with a multi-layer strategies with successive interference cancellation [61]. This configuration could be appropriate in multi-user scenarios.

- Frequency pilot tone/s : due to the proximity of certain groups of symbols for some modulation orders and the low SNR levels for which maximum shaping gains are obtained, phase noise compensation poses great difficulty. Frequency pilot tones permit blind recovery while preserving the generation structure and the receiver simplicity [46].

Chapter 3

Metropolitan networks

Metropolitan-area networks (MANs) are the bridge between access and long-haul systems. Comprehending transmission reaches from ≈100 km to ≈1000 km, this network segment gives service to large metropolitan areas, providing direct interconnection among cities, regions and even neighboring countries. Mainly due to the increasingly significant role of content-delivery networks[1], MANs' traffic is growing twice as fast as LH's and currently accounting for more than 50% of the global IP transport volume [12]. This proliferation combines with an exceptional diversity of flows to motivate the use of (i) high-speed flexible WDM transceivers, then enabling on-demand spectral allocation allowing for bandwidth savings and dynamic service support; and (ii) the use of all-optical add-drop multiplexers (OADMs), facilitating periodic spectral defragmentation, and providing transparent routing/switching capabilities that reduce the number of required transponders. As the ever-evolving connectivity and throughput requirements shape MANs infrastructure from 100-Gbps channels and point-to-point WDM rings towards ≥400-Gbps super-channels and all-optical reconfigurable mesh architectures, these two technologies rise as

[1]Content-delivery networks or CDNs are systems of distributed datacenters that deliver content to the end-user based on geographical proximity. Two distances are considered: (i) end-user to the CDN's point-of-presence, and (ii) the CDN's point-of-presence to the original server.

CDNs' datacenters dynamically cache/store large amounts of data form various origins (e.g. web and downloadable objects, applications, video streaming, social networks and any kind of cloud service) and physically bring it closer to the end-user thus improving the availability, security, and retrieval latency while offloading traffic from core backbones. Therefore, it is apparent that the larger the network spreads the better the performance will be, leading to CDNs with up to several thousands of interconnected datacenters whose traffic is inherently loaded onto metro reach.

the pillars of future metro networking [13–16].

From the implementation perspective, high-speed adaptive and multi-degree OADM-based MANs pose various important challenges. On one hand, photonic-layer networking leads to longer regenerator-free transmission distances, then worsening the end-to-end performance due to (i) filtering effects[2] accentuated by high-order OADM cascadability (up to ≈10 nodes), (ii) larger PMD[3], and (iii) OSNR degradation. On the other hand, sustainable increase of the per-channel bitrates requires using AMF to boast the SE as a means of sidestepping the economical and technological challenges that high-baudrate signaling entails. Unfortunately, AMF aggravate the OSNR sensitivity and reduce the system's tolerance to filtering effects and phase noise, hence clearly harmed by long-reach all-optical routing. All of these challenges strongly encourage the use of coherent technology supported with advanced DSP, suggesting the direct application of LH's transceivers as first approach. Nevertheless, the high (and growing) number of end-nodes in MANs imposes severe requisites in cost, footprint, and power consumption; making such straight integration prohibitive, and underscoring the relevance of low-noise optical amplification, low-attenuation large-effective-area fibers, intelligent optical spectrum managing (resource sharing), and an efficient AMF-FEC-DSP combo. These measures are common factor in high-speed MAN demonstrations (see Section 3.1), where, unlike in point-to-point (P2P) networks, the most-effective techniques for nonlinear impairment compensation are inapplicable from both the conceptual and the complexity-cost standpoint[4]. Whereas renovating the network's amplification scheme and/or fiber type boast the performance with total certainty, modulation format selection/design and bandwidth use optimization are both non-trivial actions tightly linked to the techno-economical requirements of the network. In this regard, sought solutions for MAN must

[2]Current wavelength-selective technology is unable to manage 12.5-GHz granularity without inferring notable crosstalk and spectral distortion of the signal of interest, hence requiring the use of frequency guard-bands (≈18 GHz [18]) that partly dispel the spectral savings [13, 31].

[3]The minimum tolerance is estimated around 25-30 ps for MANs [13]

[4]Approaches for nonlinearity compensation like digital back-propagation [37] and variants, inverse Volterra series [43], phase-conjugated twin-waves [44], or mid-span phase conjugation [45], rely on the exact knowledge of the physical properties of the propagation medium; which, furthermore, should remain constant from transmitter to receiver end (ideally). This fact prevents the use of those techniques to improve the communication performance across present and specially future MANs, showing many-node mesh topology and dynamic signal routing. Other methods such as co-propagating frequency tone conjugation [46] or sophisticated FEC schemes [89] may still be used.

feature [13]:

- *Compact spectral occupancy :* given a target bitrate, increasing the SE per carrier reduces the cost per bit, allows for higher degree of spectrum sharing, and reduces the crosstalk penalty from/to neighboring WDM channels. Pulse-shaping and AMF are key techniques within this group.

- *Robustness to OADM (filtering effects) :* the non-rectangular filter shape that realistic OADMs exhibit causes frequency-selective distortions whose severity aggravates as the number of traversed nodes increases. Signal pre-distortion together with spectrally efficient techniques can considerably increase the cascadability tolerance [90]. Of course, concomitant improvements in the OADMs' flatness and roll-off sharpness are most desirable.

- *Spectral granularity :* super-channels (WDM channels conformed by >1 lower-baudrate optical sub-carrier) offer enhanced intra-channel spectral flexibility at the expense higher number of low-speed electro-optical components. An intelligent exploitation of such flexibility shall make possible the use of finer-granularity WDM grids (see [18]) as compared to the analogous single-carrier configuration, thereby leading to higher SE through the reduction of idle optical spectrum. Nonetheless, it is important to realize that only WDM channels/super-channels are liable to optical routing (not the constituent sub-carriers), meaning that very broad super-channels pose difficulties in terms of network managing for spectrum sharing, signal protection and optical re/routing [13]. For its part, the lowerbound for the spectral occupancy is primarily established by the selectivity of the optical filters[5], where 37.5-GHz and 75-GHz slot granularity represent the next milestones in the managing capabilities of OADM technology.

  Further flexibility can be gained by digitally splitting the optical sub-carriers themselves into multiple low-baudrate signals. The ensuing high-selective spectral control enables accurate and dynamic signal pre-distortion, thus notably improving the SE per route, and the resiliency to filtering effects. These ultra-fine granularity spectral components are often separated and processed (in parallel) in the

[5] We define filter selectivity as the ratio between the passband width and the transition-band width.

digital domain, hence requiring additional power and computational capability.

- *Adhere to standard frequency grids :* complying with standardized frequency grids makes possible seamless (transparent) connectivity across manyfold networks (e.g. MANs and access networks). This shall permit unifying the control plane for more efficient resource allocation (including hardware management/sharing) [91].

Consequence of a balanced accommodation of these various requirements, the implementation complexity, and the throughput demand, equipment suppliers have converged to 2×sub-carrier 400-Gbps super-channels as the next evolutive step in high-speed WDM systems for MANs. This solution relies on quadrature multiplexing, polarization multiplexing (Pol-Mux), and spectral shaping to deliver 5.33 bit/s/Hz of net SE (WDM operation on the 75-GHz grid) with a maximum transmission distance of ≈500-km over uncompensated SSMF fiber link with EDFA-only amplification (longer transmission requires improvements in the noise figure of the optical amplification stages and/or the utilization of optical fibers with very low attenuation and large effective area to mitigate nonlinearities). At present, this configuration is deemed the optimum trade-off for attaining MAN's throughput-time-distance products without compromising the OADMs' selectivity, and keeping the electronics within acceptable margins of cost and stress[6] [92]. As far as the modulation format per sub-carrier is concerned, Nyquist pulse-shaped QAM has been commonly adopted. With typical roll-off factors (β)[7] ranging from 0.2 to 0.01, this option achieves notable inter-symbol interference (ISI)-free spectral compactness; favoring the spectral sharing (more nodes/transceivers can be allocated within the same network), and improving the routing/switching manageability (signals' spectrum can be readily accommodated within different standardized WDM grids). Although desirable, it is precisely the spectral stiffness that Nyquist pulse-shaped signals exhibit what prevents effective adaptability to the channel response, making this modulation approach particularly sensitive to filtering effects. Despite this issue has been so far circumvented with

[6]Other alternatives include 1×400-Gbps net rate with ≈100-GHz optical bandwidth occupancy, where small footprint and low cost are the main values; and 3×133-Gbps net-rate with ≈50-GHz optical bandwidth occupancy, focusing SE maximization.

[7]The roll-off factor is a measure of the excess bandwidth of the filter beyond the Nyquist bandwidth (Symbol rate/2). Accordingly, the total signal bandwidth is expressed as Signal bandwidth = (Symbol rate/2)$(1 + \beta)$

pre-distortion techniques and advanced DSP/FEC, such solution shows insufficient as MANs grow in size and traffic volume; thus increasing the average number of intermediate nodes per route, and aggravating the over-filtering per node due to the compression of the WDM spectrum together with the deficient selectivity of current OADM technology for fine granularity ($\leq$37.5 GHz). This situation has engaged the attention of the research force, resulting in diverse proposals in which further super-channel partitioning (<32-GBd sub-carriers) is contemplated with the main purpose of enabling more precise spectral shaping and lower implementation penalties [15, 93]; nevertheless, at the expense of abundant optical parallelism, definitely contraposing the aspirations for inexpensive and small-footprint transceivers.

In this regard, and supported by the improvements in the speed and resolution of digital-analog interfaces, digital sub-carrier de/multiplexing is increasingly looked into for alleviating the mentioned trade-off and providing flexible software-defined control of the sub-carriers' properties [94]. This is the context of the next investigation, where we report on a novel DSP-mediate method for signal multiplexing and modulation called coherent multi-band carrierless amplitude/phase modulation (C-MultiCAP). Unlike in other techniques such as OFDM [95], the ensemble spectral shape and the total bitrate in C-MultiCAP are tailored in a sub-band by sub-band basis, enabling individual adjustment of the central frequency, baudrate, SE, as well as number of dimensions per sub-band and per polarization. As a consequence of these numerous degrees of freedom, unmatched flexibility-robustness ratio is achieved without compromising the SE at the expense of processing complexity. Resulting from an adaptation of the original work by M. I. Olmedo et al. [96] for operation in coherent communications, we propose and experimentally demonstrate for the first time C-MultiCAP in fiber-optic communications as potential multiplexing plus modulation technique for future MAN's transceivers.

This chapter is subdivided into four blocks. First, the most relevant research work in the context of high-speed WDM MANs is presented, making explicit distinction between single-carrier and super-channel approaches, as well as among diverse link configurations. This information bases the subsequent positioning of the present investigation. In the second block, the essentials of single- and multi-band multi-dimensional carrierless modulation are thoroughly discussed as the fundamental building blocks in C-MultiCAP. Signal properties and practical implications of using this

multiplexing technique are the main focus. In the last part of this second block, C-MultiCAP detection is elaborated from the DSP perspective. The core processes for enabling successful equalization and demodulation are explained in detail, together with the constraints and assumptions that these imply. In the third block, the experimental proof-of-concept of a C-MultiCAP signal successfully transmitted over a MAN-reach scenario is introduced. Finally, the work is summarized, comprehending a motivation recap, main results and conclusions, and plausible future work lines.

3.1 Work positioning

In this section, the most relevant ≥4-bit/s/Hz net SE transmission demonstrations are presented in the context of WDM MANs. For concision and taking into consideration MANs' constraints and operational requirements, we restrict the covered literature to >100-Gbps net rate WDM channel/super-channels with ≤175 GHz of spectral occupancy, standardized WDM frequency grid/slots [18], and employing compatible techniques for impairment compensation (see footnote 4). The results are organized into two subsections according the number of sub-carriers constituting the WDM channel, that is single-carrier and multiple sub-carriers (super-channels). In addition, further subdivision is realized within each category based on the span configuration, where the usual combination of SSMF or non-zero dispersion-shifted fiber (NZDSF) plus EDFA is distinguished from any other alternative. In the end, this investigation is contextualized and the main contributions highlighted.

Single-carrier

High-baudrate single-carrier approaches target the minimization of the transceivers' footprint by reducing the number of electro-optical components and thereby the associated cost. On the other hand, the ensuing broad spectrum often undergoes acute over-filtering (even in the absence of OADMs), then requiring the use of sophisticated DSP, powerful FEC and/or pre-emphasis techniques that entail extra computational load and power consumption. Recent investigations include:

SSMF/NZDSF plus EDFA :

- [97–99]: 1 WDM channel modulated with DP 56-GBd 16-ary QAM is transmitted over 12×75-km and 16×75-km uncompensated SSMF

spans (900 km and 1200 km in total respectively) with EDFA-based amplification. Optical pulse-shaping with $\beta = 0.2$ suffices for fitting the signal within one 100-GHz grid slot, yielding 4.48 bit/s/Hz gross SE per channel. In order to counteract the severe ISI, the DSP at the receiver features 3-, 5-, or 7-symbol memory MAP demapping. The *a priori* estimations are calculated during system's initialization (using training sequence) and subsequently stored in a look-up table to minimize the computational load. Standard hard-decision 7%-overhead Reed-Solomon FEC is considered [20]. This work outperforms preceding investigations for single-carrier DP 56-GBd 16-ary QAM in terms of throughput-times-distance product [100–102].

- [103]: 5 WDM channels implementing DP 42.66-GBd 64-ary QAM each are transmitted over 2×80-km uncompensated SSMF spans (160 km in total) and 3×100-km uncompensated ULAF spans (300 km in total), both cases employing EDFA-only amplification stages. Digital pulse-shaping with $\beta = 0.15$ (1.5 S/symbol) and pre-emphasis are employed, allowing for 50-GHz frequency-slot granularity (one slot per channel with ≈10.24 bit/s/Hz gross SE) while keeping the linear cross-talk within negligible margins. Soft-decision 24%-overhead spatially-coupled LDPC convolution codes are used for FEC ($\approx 5 \cdot 10^{-2}$ pre-FEC BER threshold).

Other span configurations :

- [104] and the follow-up [105]: 1 WDM channel modulated with DP 124-GBd 32-ary QAM is transmitted over 12×55-km uncompensated low-loss ULAF spans (660 km in total) using EDFA-only amplification. Four-fold spectral slicing is employed at the transmitter to synthesize the high-baudrate signal, thus relaxing the electro-optical bandwidth requirements and enabling near-Nyquist digital pulse-shaping to fit the signal in a single 125-GHz frequency slot (9.92 bit/s/Hz gross SE per channel). After detection with a single coherent receiver, the signal is post-processed offline. Powerful soft-decision spatially-coupled LDPC convolutional FEC code with 24% overhead is implemented within standard BICM setting ($\approx 4.5 \cdot 10^{-2}$ pre-FEC BER threshold).

- [106]: 5 WDM channel modulated with DP 72-GBd 64-ary QAM is transmitted over 4×100-km uncompensated ULAF spans (400 km

in total) using hybrid Raman/EDFA amplification stages. The signal generation is performed with two differential-output DACs at 1 S/symbol (72 GS/s) per channel, and one programmable optical filter is employed to fit each channel in one 100-GHz frequency slot (8.64 bit/s/Hz gross SE per channel) and pre-emphasize the spectrum to compensate for bandwidth limitations. An adaptive real-valued 8×4 multi-input multi-output (MIMO) (40-symbol memory) feed-forward equalization accounting for differences in the linear response between quadrature components is implemented. Hard-decision FEC with ≈35% redundancy is considered.

- [107]: 20 WDM channels modulated with DP electrical time-division multiplexing (ETDM) 43-GBd 64-ary QAM are transmitted over 12×50-km uncompensated low-loss ULAF spans (600 km in total) using EDFA-only amplification. Optical pre-emphasis is carried out to alleviate filtering impairments while fitting the channels in the 50-GHz grid (≈10.32 bit/s/Hz gross SE). In order to compensate for the ISI, triple stage equalization and MAP detector are implemented. Soft-decision 20%-overhead quasi-cyclic LDPC is employed as reference FEC code [108].

- [109]: 1 WDM channel modulated with DP Nyquist optical time-division multiplexing (OTDM) 160-GBd 64-ary QAM is transmitted over 2 spans constituted by 50-km super-large-area fiber (SLAF) plus 25-km inverse dispersion fiber (IDF) (150 km in total) using EDFA-only amplification. The advantage of using Nyquist pulses as compared to other shapes is two-fold: fit the signal within the 175-GHz grid (≈10.97 bit/s/Hz gross SE) and exploit their time-domain orthogonality to facilitate pulse demultiplexing. Hard-decision 7% overhead FEC is considered.

Multiple sub-carriers

These approaches increase the number of electro-optical components in favor of routing flexibility, enhanced signal protection/robustness against filtering impairments, and lower stress on the electronics; hence potentially enabling higher throughput-times-distance products owed to the minimization of the implementation penalty. In addition, the milder bandwidth requirements allow for using commonplace DACs for signal generation, making possible near-Nyquist spectral compactness via digital pulse-shaping

with arbitrary-low $\beta > 0$; and the use of sophisticated modulation plus multiplexing schemes (e.g. OFDM). Investigations in controlled laboratory environment include:

SSMF/NZDSF plus EDFA :

- [110]: 3 WDM super-channels constituted by 5×28-GBd pre-filtered return to zero (RZ) (50% duty-cycle) DP 16-ary QAM (1.12 Tbps gross rate per super-channel) are transmitted over 400-km uncompensated SSMF link with a total of 2 OADM passes using EDFA-only amplification. Each super-channel is fit in one 162.5-GHz frequency-slot, delivering ≈6.89 bit/s/Hz gross SE per super-channel. The DSP in the receiver side comprises duobinary filtering plus maximum-likelihood sequence estimation (MLSE) for boasting the SNR. Hard-decision 7%-overhead Reed Solomon code is assumed for FEC [20].

- [111]: 1 WDM super-channel constituted by 2×32-GBd DP 16-ary QAM sub-carriers is transmitted in real time along with another 58×25-GBd 50-GHz spaced DP-QPSK neighboring WDM channels over 10×100-km uncompensated SSMF spans (1000 km in total), and 5×100-km uncompensated NZDSF spans (500 km in total) using only EDFAs in both scenarios. The 200-Gbps sub-carriers are spaced by 37.5 GHz after digital Nyquist pulse-shaping, and both fit in 75-GHz optical bandwidth (≈6.82 bit/s/Hz gross SE per super-channel). Soft-decision 20%-overhead quasi-cyclic LDPC is employed as reference FEC code [108].

- [112]: investigation of optimum sub-carrier power and frequency guard-band between super-channels to maximize SE and reach in optically-routed flexible networks. Numerical study is performed for terabit-class super-channel comprising 5×32-GBd sub-carriers modulated with DP 16-ary QAM (1280 Gbps gross rate per super-channel) transmitted over 15×60-km uncompensated SSMF spans (900 km in total) and EDFA-based amplification. Digital Nyquist pulse-shaping with $\beta = 0.15$ is employed, permitting 35-GHz spacing between sub-carriers and 175-GHz optical bandwidth occupancy per super-channel. The results prove 21% reach improvement for as low as 12.5 GHz guard-bands, and up to 62% increase in SE-times-reach product for full optimization.

Other span configurations :

- [113]: 1 WDM super-channel constituted by 2×28-GBd sub-carriers using pre-filtered RZ (50% duty-cycle) DP 16-ary QAM (448 Gbps gross rate per super-channel) is recirculated 10 times through a fiber-loop made up of 1×72-km pure silica core fiber (PSCF) (720-km total propagation distance) and one OADM; Raman amplification with counterpropagating pump plus EDFA are employed for attenuation compensation. The use of pre-filtered RZ pulse-shaping enabled the use of 75-GHz slot granularity (5.97 bit/s/Hz gross SE per super-channel) while achieving high resilience to filtering effects without either digital pre-emphasis or many-tap finite impulse response (FIR) Nyquist filters. Standard hard-decision 7%-overhead Reed-Solomon FEC is considered [20].

  With the focus on inter-channel interference analysis, the same group extends [113] by demonstrating 5 WDM super-channel transmission over 3 turns on a recirculating loop comprising 4×50-km plus 1×26-km PSCF (678 km in total) and one OADM using EDFA-only amplification [114]. Each super-channel comprehends 2 sub-carriers modulated with pre-filtered RZ (50% duty-cycle) DP 16-ary QAM, allowing for 75-GHz of optical bandwidth occupancy per super-channel (same configuration as in [113]) and 35-GHz spacing between all neighboring sub-carriers.

- [115]: 4 WDM super-channels constituted by 2×32-GBd DP 16-ary QAM sub-carriers (512 Gbps gross rate per super-channel) are transmitted over 3×210-km uncompensated SSMF spans (630 km in total) and amplified with hybrid stages comprising bidirectional Raman pumping and one EDFA. Digital pulse-shaping enables the use of 37.5-GHz frequency-slot granularity (one slot per sub-carrier), thus leading to ≈6.82 bit/s/Hz gross SE per super-channel. Soft-decision 24.5%-overhead spatially-coupled irregular LDPC code is concatenated with an outer BCH de/coding block (25.5% of total redundancy) for FEC [116].

- [117]: 3 WDM super-channels constituted by 3×46-GBd DP 32-ary QAM sub-carriers (1.38 Tbps gross rate per super-channel) are transmitted over uncompensated SSMF with hybrid Raman-EDFA opti-

cal repeaters for attenuation compensation. The three optical sub-carriers are packed in 140-GHz (allowing for 150-GHz frequency grid including 1-GHz guard-bands) through digital Nyquist pulse-shaping with $\beta = 0.1$, yielding a gross SE per super-channel of ≈9.2 bit/s/Hz. The maximum transmission reach is evaluated with and without intermediate OADMs, showing successful signal recovery after 600-km of propagation with 3 OADM passes and up to 800 km without filters. Soft-decision 32%-overhead convolutional LDPC code is considered for FEC, enabling error-free demodulation at BER$\leq 5.3 \cdot 10^{-2}$.

- [118]: 5 WDM super-channels constituted by 5×9-GBd DP 32-ary QAM sub-carriers (450 Gbps gross rate per super-channel) are transmitted over 8×100-km ULAF (800 km in total) using Raman amplification plus EDFA. Near-Nyquist digital pulse-shaping ($\beta = 0.01$) permits fitting the signal within just 45.8 GHz optical bandwidth (50-GHz WDM grid), which gives ≈9 bit/s/Hz gross SE per super-channel. Optical pre-emphasis is used to counteract the filtering effect of a single OADM in the transmitter, thereby achieving 100% longer reach as compared to the case without pre-compensation. Hard-decision continuously interleaved BCH FEC code with 7% redundancy is contemplated [119].

- [120]: 5 WDM super-channels carrying 2×9-GBd DP 64-ary QAM sub-carriers interleaved with 3×9-GBd DP hybrid 64-32-ary QAM sub-carriers (504 Gbps gross rate per super-channel) are transmitted over 12×100-km uncompensated ULAF spans (1200 km in total) using all-Raman amplification. Band-limiting effects are alleviated via frequency-domain digital pre-distortion and Nyquist pulse-shaping with $\beta = 0.01$. Operation in the 50-GHz frequency grid is demonstrated (≈45.8-GHz optical bandwidth occupancy), resulting in ≈10.1 bit/s/Hz gross SE per super-channel. The use of hybrid QAM is key to permit successful post-FEC demodulation in the considered scenario, where the soft-decision 20%-overhead quasi-cyclic LDPC code is employed as reference [108].

- [121]: 90 WDM super-channels constituted by 8×5.6-GBd DP 64-ary QAM sub-carriers (538 Gbps gross rate per super-channel) are transmitted over 3×80-km uncompensated PSCF spans (240 km in total) using Raman amplification. The super-channel is fit within one 50-GHz frequency slot via digital Nyquist pulse-shaping with $\beta = 0.1$,

providing ≈10.76 bit/s/Hz gross SE per super-channel. A radio-frequency pilot tone is digitally inserted in the transmitter side for enabling robust and data-independent phase noise compensation, allowing for stable performance for up to ≈15 dBm total launch power. Hard-decision continuously interleaved BCH FEC code with 7% redundancy is contemplated [119].

- [122]: 370 WDM super-channels constituted by 4×6-GHz DP OFDM sub-carriers modulated with 128-ary QAM (294 Gbps net rate per super-channel) are transmitted over 3×55-km uncompensated SSMF (165 km in total) using all-Raman amplification. Frequency locking among the sub-carriers permits using down to 1-GHz frequency guard band without performance degradation and thus fitting the super-channel in one 25-GHz slot (≈11.76 bit/s/Hz gross SE per super-channel). The DSP comprehends standard OFDM TX/RX blocks with the exception of a dual-stage phase noise compensation algorithm in the receiver side, which exploits the residual direct current (DC) component that appears after modulation to realize coarse estimation followed by fine tracking employing pilot sub-carriers. Hard-decision 7%-overhead Reed-Solomon FEC is considered [20].

- [123]: 1 WDM super-channel constituted by 12×3.94-GHz DP OFDM sub-carriers using offset 64-ary QAM for the payload (516.17 Gbps gross rate per super-channel) is transmitted over 5×80-km uncompensated SSMF spans (400 km in total) using Raman-only amplification. The use of offset-QAM technique in combination with OFDM achieves remarkable side-lobe suppression, thus providing a quasi-rectangular spectral shape with high confinement (≈49.8-GHz total optical bandwidth occupancy) that avoids the need for timing or frequency alignment among sub-carriers (OFDM sub-bands) to form the super-channel. Considering 50-GHz frequency-slot granularity, the net and gross SE per super-channel equals ≈8.6 bit/s/Hz and ≈10.32 bit/s/Hz respectively.

As far as field-trial demonstrations are concerned:

- [124]: 2 WDM super-channels constituted by 2×32.5-GBd DP 16-ary QAM sub-carriers (520 Gbps gross rate per super-channel) are transmitted in real time over a legacy multi-vendor 400-km uncompensated SSMF link with EDFA-only amplification in Turin area,

Italy. Each sub-carrier is electronically Nyquist pulse-shaped to fit in one 50-GHz frequency slot, yielding 100 GHz per 400-Gbps super-channel (4 bit/s/Hz net SE and 5.2 bit/s/Hz gross SE). Stable and successful post-FEC demodulation is reported over >65 hours, where the FEC code is composed of an inner soft-decision LDPC convolutional block and an outer hard-decision Reed-Solomon block for a total 23% redundancy.

- [125]: 1 WDM super-channel constituted by 2×32.5-GBd DP 16-ary QAM sub-carriers (520 Gbps gross rate per super-channel) is transmitted in real time over a commercial 550-km uncompensated SSMF link with EDFA-only amplification between Paris and Lyon, France. Uninterrupted post-FEC error-free demodulation is reported in both directions for 24 hours with more than 2 dB average margin in Q^2-factor at 193.1 THz center frequency. Soft-decision FEC is employed.

Deprived of intermediate electronic signal regeneration and end-to-end nonlinearity compensation, research activities within MAN context consistently evince the difficulty of attaining ≥5.33 b/s/Hz net SE over >500 km of SSMF/NZDSF and EDFA amplification while complying with standard ≈2-dB OSNR industrial margins. This fact holds even without mid-link OADM passes[8], becoming further challenging in large optically-switched networks carrying high-baudrate signals, where filtering distortions predictably dominate the performance [13, 90]. Aiming for boosting the receiver sensitivity irrespective of the node-cascading degree, numerous studies opt to improve the link components, that is, amplification stages with lower noise figure and/or fibers with lower attenuation-per-kilometer and larger effective area. Albeit net SE-times-reach products as high as 9600 bit/s/Hz-km have been proved in such scenarios [120], network operators rather deem this potential upgrade as an undesirable concern in terms of cost and deployment time, prioritizing the reuse of the existing equipment to the largest possible extent [111]. All of the mentioned motivates the development of de/multiplexing and de/modulation approaches focusing on maximizing the receiver sensitivity at the same time that providing high spectral efficiency and robustness to filtering effects; which partly justifies

[8]Note that all cited investigations with the exception of [113, 114] employ WDM technology in P2P configuration with ≤2 OADMs per path.

the abundant research on WDM super-channels and alternatives to raised-cosine pulse shapes (e.g. RZ [110,114] or doubinary [126]). With such target and considering modulation schemes, OFDM represents the most resorted alternative to the rigid QAM in order to equip the transceivers with accurate digital control on the sub-carriers' signal properties; featuring down to sub-gigahertz granularity spectral control and rate adaptability, while delivering near-Nyquist SE. Regrettably, because OFDM's functioning relies on the frequency-wise orthogonality among typically hundreds of superimposed low-baudrate modulated sinusoids, it presents high sensitivity to effects causing inter-carrier interference (ICI) and phase noise, as well as it exhibits severe peak-to-average power ratio (PAPR) that increases the quantization noise and makes it susceptible to nonlinearities.

It is thus clear that the selection/design of a modulation technique that meets the requirements of future MANs still eludes a satisfactory solution as of today, where the usually considered alternatives (i.e. QAM and OFDM) show a series of complementary pros and cons that suggest looking into approaches offering an intermediate compromise. That is the purpose of C-MultiCAP, a flexible multiplexing plus modulation technique conceived for coherent communications, that leverages on DSP in the transmitter and receiver sides to sub-divide the electrical bandwidth into a user-defined number N-dimensional pass-band signals. Besides supporting multi-dimensional modulation format per-band, their individual center frequency, baudrate and excess bandwidth (e.g. determined by β) are also liable to adjustment, further allowing for near-Nyquist spectral compactness and improved robustness and sensitivity through vast spectral control.

In the following, we develop in detail on the generation and detection of C-MultiCAP prior to discuss the results of the first experimental MAN-reach transmission demonstration. These results show successful signal recovery after >200-km and >400-km SSMF EDFA-only link for 300-Gbps and 200-Gbps net rate respectively employing standard 25G-class electro-optical components.

Contributions of this work

- Original idea: development of C-MultiCAP for enabling the advantages of carrierless amplitude/phase modulation (CAP) and multi-band carrierless amplitude/phase modulation (MultiCAP) in coherent communications. Development of the RX DSP structure and the multi-dimensional parallel equalization approach that makes possible impairment compensation and correct signal recovery. Proposal of

C-MultiCAP for its use in optically-routed MANs.

- Experimental results: experimental demonstration of C-MultiCAP in metro-reach coherent optical communications.
- Related first-author publications: [127] (conference).

3.2 Multi-band carrierless amplitude/phase modulation for coherent communications (C-MultiCAP)

C-MultiCAP results from the evolutionary adaptation of CAP, a pass-band 2D modulation approach introduced by D.D. Falconer in 1975 [128] and later generalized to multiple dimensions; firstly to overcome the original susceptibility to filtering distortions and relax the stress on the electronics when digitally generated (MultiCAP, [96]); and finally making possible its use in coherent communications to boost throughput and reach [127].

In this section, C-MultiCAP generation and detection are discussed and contrasted with the predecessors', paying special attention to the influence of the signal properties on the general performance, as well as the associated challenges from the implementation standpoint.

3.2.1 Generation and signal properties

The exceptional property in all CAP flavors is the ability to de/multiplex $\geq$2 independent data streams within a given bandwidth while maintaining end-to-end system linearity, that is, FIR filters suffices for multi-dimensional signal generation and data recovery in the absence of distortions. With such filters fundamentally determining all the signal properties, understanding the restrictions and variables governing their design is crucial to assure correct adaptation to the available resources and channel properties for maximum sensitivity and SE. This is the issue covered in the following subsection.

Starting from it most basic form, CAP, the mathematical foundations in linear multi-dimensional carrierless de/multiplexing are presented, as well as the practical value and potential applications of the resulting signal properties. Following the initial observations, the generalization to multiple bands is briefly discussed as a bridge towards C-MultiCAP, whose characteristics are detailed along with the technical requisites to permit successful demodulation.

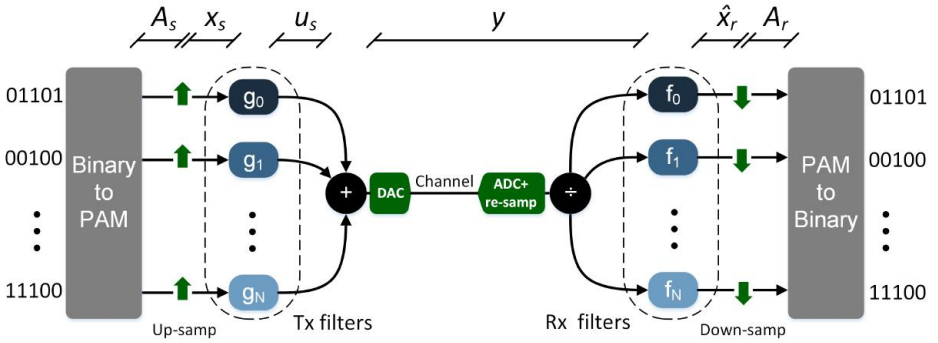


Figure 3.1: General schematic diagram of CAP-based systems. Transmitter on the left-most side and receiver on the right-most side. The mathematical nomenclature of the signal at different stages is included on the top part.

3.2.1.1 CAP

Figure 3.1 shows the schematic diagram illustrating the general principle for CAP-type signal generation and reception. In the transmitter side, N parallel streams of binary data[9] are modulated in amplitude prior to their convolution with respective filters, g_s, where $\{s \in \mathbb{N}; s \leq N\}$. The output from these filters is linearly superimposed and propagated over the channel. The received waveform is N-fold split/copied and processed individually with another set of N filters, f_r; followed by regular amplitude-to-binary demapping according to the same labeling strategy as the transmitters'. Mathematically, the noise-free characteristic equation from the input of transmitters' filters, x_s, to the output of receivers', $\hat{x}_r$, can be expressed as (see Figure 3.1 for full nomenclature):

$$\hat{x}_r[n] = (f_r * y)[n] = f_r[n] * \left(\sum_{s=1}^{N} u_s[n]\right) = f_r[n] * \left(\sum_{s=1}^{N} g_s[n] * x_s[n]\right) \tag{3.1}$$

with $x_s[n]$ defined as:

$$x_s[n] = \begin{cases} A_s[k] & \text{for } n = kT \text{ with } k \in \mathbb{Z} \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

[9]For concision, the transition from serial source-coded binary to parallel and channel-coded binary is assumed and thus not depicted.

where T is the symbol period in samples (i.e. over-sampling factor) and A_s is a real value that stands for the amplitude at symbol instants. Notice that different mapping alphabets can be used for each branch (e.g. A_0 and A_3 may be obtained from 4-ary and 8-ary pulse-amplitude modulation (PAM) respectively).

Then expanding the convolutions in 3.1 and considering 3.2, we obtain the general expression governing the system depicted in Figure 3.1:

$$\begin{aligned}\hat{x}_r[n] &= \sum_{m=-\infty}^{\infty} f_r[m] \left(\sum_{s=1}^{N} \sum_{l=-\infty}^{\infty} g_s[l] x_s[n-m-l] \right) \\ &= \sum_{m,l=-\infty}^{\infty} \sum_{s=1}^{N} f_r[m] g_s[l] x_s[n-m-l] \\ &= \sum_{s=1}^{N} x_s[kT] \sum_{m=-\infty}^{\infty} f_r[m] g_s[n-kT-m] \\ &= \sum_{s=1}^{N} A_s[k] \left(f_r * g_s\right)[n-kT]\end{aligned} \tag{3.3}$$

whose impulse response $h_{rs}[n]$ (see[10]), defined in a MIMO system as the collection of impulse responses for each input channel , and the corresponding frequency response $H_{rs}(f)$ are given by:

$$h_{rs}[n] = \left(f_r * g_s\right)[n], \tag{3.4}$$

and

$$H_{rs}(f) = F_r(f) G_s(f) \tag{3.5}$$

Now, by constraining 3.4 and 3.5 based on the desired functional qualities of the system in time- and frequency-domain, the implementation

[10]Because linear MIMO systems are prone to matrix formulation, the matrix-form equation equivalent to 3.3 is provided below for the readers' convenience:

$$\hat{\mathbf{x}} = \mathbf{F} * \mathbf{G} * \mathbf{x} = \mathbf{H} * \mathbf{x},$$

where $\mathbf{H} = \mathbf{F} * \mathbf{G} = h_{rs}[n]$ is the $N \times N$ matrix describing the input-output response, with N equal to the number of transmitted parallel data streams and $\mathbf{F} = \{f_1[n], f_2[n], ..., f_N[n]\}$ and $\mathbf{G} = \{g_1[n], g_2[n], ..., g_N[n]\}^\top$. For their part, $\mathbf{x} = \{x_1[n], x_2[n], ..., x_N[n]\}$ and $\hat{\mathbf{x}} = \{\hat{x}_1[n], \hat{x}_2[n], ..., \hat{x}_N[n]\}$ represent the instantaneous transmitted and received real-valued signals respectively.

conditions on the TX/RX filters (g_s, f_r) can be extrapolated. Next we particularize for CAP signals' properties as originally described in [128] and later thoroughly detailed in [129], that is, a linear ISI-free multi-dimensional modulation and multiplexing technique delivering optimum end-to-end SNR sensitivity in the presence of AWGN.

Time-domain constraints

ISI-free communication : zero end-to-end ISI avoids the need for equalization or reduces its complexity. In the system described by 3.3, the ISI-free condition is given by:

$$h_{rs}[n] = \begin{cases} 1 & \text{for } n = 0 \\ 0 & \text{for } n = kT \text{ with } \{k \in \mathbb{Z}; k \neq 0\} \\ \text{undefined} & \text{otherwise} \end{cases}, \tag{3.6}$$

Accordingly, 3.3 results in $\hat{x}_r[kT] = \sum_{s=1}^{N} A_s[k]\ \forall r$. Note how cross-channel/dimension interference impairs the recovered signal if no further action is taken. In the frequency domain, the ISI-free condition is equivalent to (Nyquist ISI criterion [130]):

$$\sum_{j=-\infty}^{\infty} H_{rs}\left(f - \frac{j}{T}\right) = T\ \forall f \tag{3.7}$$

As an example, Figure 3.2 shows one general type of frequency response that meets 3.7. Raised-cosine pulse shape belongs to this class.

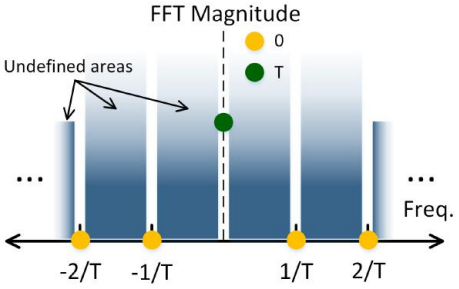


Figure 3.2: Spectral illustration of a general system response showing zero end-to-end ISI.

Multi-dimensionality : this quality plus 3.6 require the impulse response describe an $N\times N$ MIMO system providing transparent communication from each of the N inputs to exclusively one of the outputs (zero cross-channel/dimension interference) at the symbol instants ($n = kT$). For instance, the most immediate configuration is $\hat{x}_r = x_s \ \forall r = s$ (i.e. $\hat{\mathbf{x}} = \mathbf{x}$ in footnote 10), which is conditional upon:

$$h_{rs}[n] = \begin{cases} 1 & \text{for } r = s \text{ and } n = 0 \\ 0 & \text{for } r \neq s \text{ and } n = 0 \\ 0 & \text{for } n = kT \text{ with } \{k \in \mathbb{Z}; k \neq 0\} \\ \text{undefined} & \text{otherwise} \end{cases} \tag{3.8}$$

The top case in 3.8 imposes transparent input-output communication at symbol instants, the second forces the necessary orthogonality between any other TX-RX filter-pair combination, and the third assures zero ISI. In particular, this solution restricts the system to a static configuration in which $\hat{x}_r = x_s$ if and only if $r = s$ (color-matched pairs in Figure 3.1); however, it should be noted that dynamic filter adjustment or interchange would allow for $N\times N$ transparent switching in whatever domain (optical or electrical) the filters are implemented. Herein, and without loss of generality, we adhere to the former for simplicity.

SNR maximization : finally, assuming AWGN channel, employing matched filters optimally maximizes the SNR in the receiver. With the general TX-RX matched-filter pair being defined as $h_{TX}[n] = h^*_{RX}[-n]$, we obtain the following constrain in our system:

$$f_r[n] = g_s[-n], \text{ or } f_r[-n] = g_s[n] \text{ for } r = s \tag{3.9}$$

Because g_s and f_r are real-valued filters $\forall r, s$, conjugation has no effect and can be neglected. Under 3.9, the impulse response 3.4 is restated as:

$$h_{rs}[n] = g_r[-n] * g_s[n] \tag{3.10}$$

In summary, 3.8 and 3.9 constitute the essential time-domain restrictions to be forced on 3.4 so as to enable CAP-type communication.

Frequency-domain constraints

As far as spectral requirements are concerned, it is highly desirable to minimize the bandwidth occupancy of our signal such that the spectral

components beyond certain frequency can be forced to zero while assuring compliance with 3.8. From the implementation perspective, this translates into higher SE and thus denser WDM packing. Mathematically:

$$|H_{rs}(f)| = \begin{cases} \text{undefined} & |f| < BW_{min} \\ 0 & \text{otherwise} \end{cases}, \tag{3.11}$$

which after 3.10:

$$|G_s(f)| = \begin{cases} \text{undefined} & |f| < BW_{min} \\ 0 & \text{otherwise} \end{cases}, \tag{3.12}$$

where BW_{min} is given by the *dimensionality theorem* [130], which is a generalization of the Nyquist–Shannon sampling condition for $N > 1$. It states that the number of independent pieces of information (dimensions, N) that can be conveyed by a band-limited waveform or communication system is proportional to the product of such bandwidth (BW_{min}) and the time allowed for transmission of the information (T):

$$BW_{min} = \frac{N}{2T} = \frac{NF_B}{2} \tag{3.13}$$

where $F_B = \frac{1}{T}$ is the symbol rate. After 3.13, 3.12 accepts the graphical representation in Figure 3.3, which depicts instances of the minimum optical bandwidth occupancy of 1-, 2-, and 3-dimensional CAP signals as a function of the symbol rate.

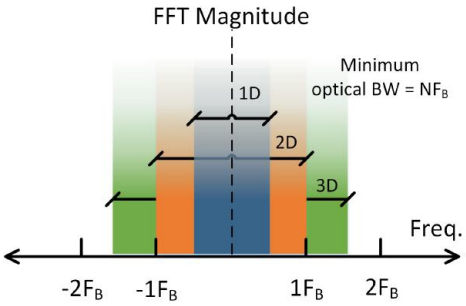


Figure 3.3: Minimum spectral occupancy of multi-dimensional CAP signals for $N \in \{1, 2, 3\}$.

One relevant observation becomes apparent from 3.13 and Figure 3.3, it is not possible to increase the SE only by adding further dimensions. Given

a fixed symbol rate, increasing the number of orthogonal filters requires a proportional increase in the signal bandwidth; or alternatively, given a fixed signal bandwidth, increasing the number of orthogonal filters requires a proportional decrease of the symbol rate. This makes the maximum achievable SE is independent of N, resulting in a SE upperbound for CAP equal to the best-case SE across all constituent one-dimensional (1D) signals. On the other hand, these independent 1D signals remain operational for multi-dimensional modulation format design, whereby the receiver sensitivity is potentially increased by inducing a controlled degree of correlation among the amplitude modulators of the employed dimensions. In this way, by distributing the symbols on the $\leq$N-dimensional constellation space to maximize their relative Euclidean distance[11], achievable SNR sensitivity gains up to 1-dB with respect to DP-QPSK have been theoretically demonstrated at BER=10^{-3} for $N = 8$ and 1 b/symbol/dimension [131, 132].

The *dimensionality theorem* admits a second interpretation of practical relevance. In this case, we generalize the expression by considering excess bandwidth:

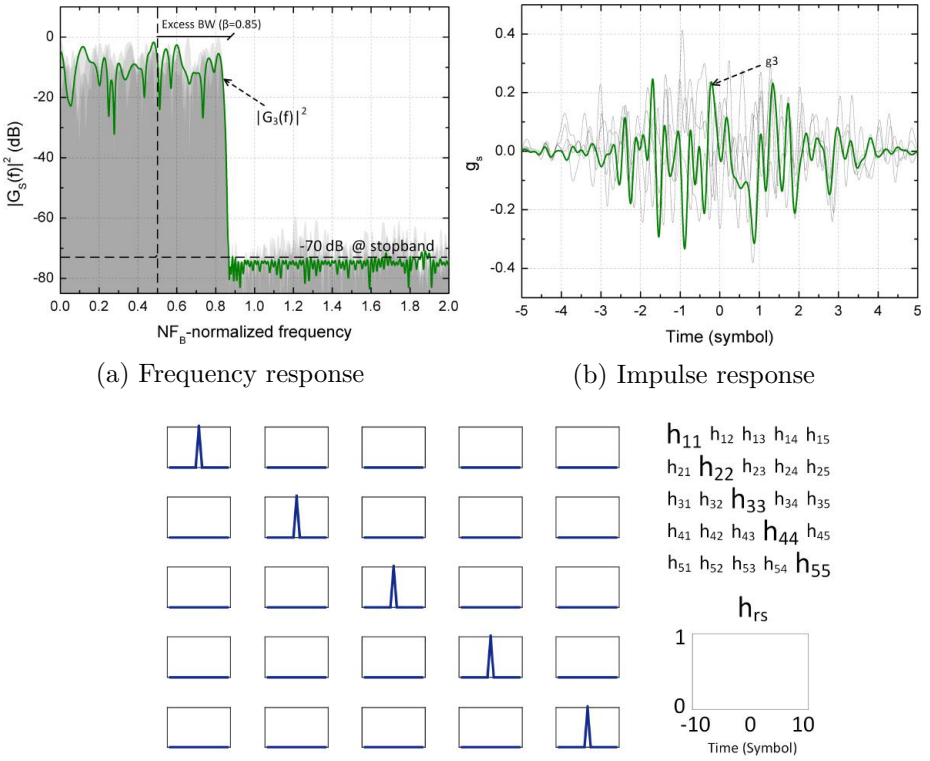
$$2BW_{min} = NF_B(1 + \beta) = F_{S;min} \tag{3.14}$$

where $F_{S;min}$ is the required sampling rate. Then normalizing by F_B, the minimum number of samples per symbol (S_{sym}) is obtained ($S_{sym} \in \mathbb{N}$ assumed):

$$S_{sym} = \lceil N(1 + \beta) \rceil \tag{3.15}$$

It becomes clear one major implementation challenge in ND-CAP system: $\geq N$ S/symbol are needed for perfect signal generation/reconstruction (no information loss). This means that, considering a fixed β, ND-CAP requires N times longer FIR filters than PAM for achieving the same SE. Over and above, ND-CAP employs N times more filters; resulting in a total of N^2 times more processing complexity than the 1D counterpart, and only for signal generation/multiplexing and detection/demultiplexing. In absolute numbers, the total amount of operations results from a compromise between the target β and the desired pulse-shaping fidelity (to avoid premature error floors). Leveraging the β-versus-tap characterization for RRC

[11] The minimum squared Euclidean distance is the optimum optimization criterion in the presence of isotropic noise (e.g. AWGN). Different transition probabilities, such as phase noise's, need consequent revision of the criteria in order to obtain optimum performance.



(a) Frequency response

(b) Impulse response

(c) System matrix $\mathbf{H} = h_{rs}[n]$.

Figure 3.4: Characteristics graphs for one solution to 5D-CAP employing *minimax* rule and constraints 3.8, 3.12, and 3.10.

filters in [96] and empirical observations after [127, 133–135], $\geq 10N$-sample FIR is required for $\beta \geq 0.1$ at BER=10^{-9}.

Filter calculation

Once the desired signal properties are mathematically characterized, the basis functions corresponding to the series of discrete CAP filters g_s meeting those requisites is calculated according to certain decision method. In this regard, the *minimax* rule has been extensively employed in the literature for this concrete purpose [136–140], aiming for the minimization of the largest value of a set of functions (e.g. the spectral components beyond BW_{min} for every filter) under a set of inequality, equality or even non-

linear constraints. As an example, Figure 3.4 shows the squared frequency response magnitude, impulse responses, and the characteristic system matrix $\mathbf{H} = h_r s[n]$ of one possible 5D-CAP solution employing *minimax*. The excess bandwidth is fixed at a relaxed 185% (β=0.85) of the minimum bandwidth $\frac{NF_B}{2}$.

Further conditions shall be imposed on the frequency response to suit the signal to particular scenario-dependent requisites, for instance, no DC component [139]:

$$|G_s(f)| = \begin{cases} \text{undefined} & 0 < f < BW_{min} \\ 0 & \text{otherwise} \end{cases}, \tag{3.16}$$

or forcing identical spectra among all N frequency responses, hence obtaining homogeneous and power equalized frequency responses [140]:

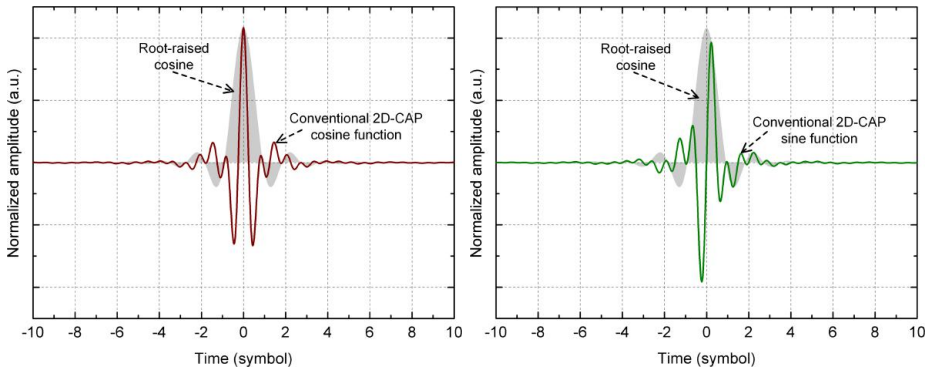
$$< \left(|G_r(f)| - |G_{r'}(f)|\right)^2 >= 0 \text{ for } \{r' \in \mathbb{Z}; r' \neq r \text{ and } r' \leq N\} \tag{3.17}$$

The reader is referred to the cited literature for graphical examples with constraints 3.16 and 3.17.

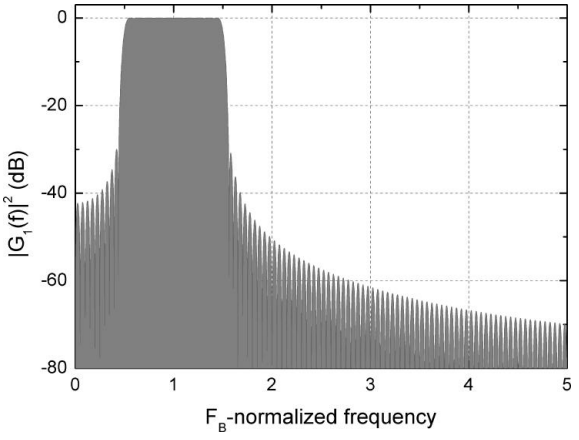
Original 2D-CAP [128, 129]: Initially, CAP was conceived as a concrete ISI-free 2D pass-band line code with adaptive bandwidth occupancy. Essentially, a variant implementation to standard QAM's in which high simplicity was comparatively accomplished by linearizing signal generation and detection without trading SE or performance. Unlike in the general approach previously discussed, in this CAP form the two orthogonal filters exhibit close-form expressions, resulting from the combination of two powerful tools: (root) raised-cosine filter type ($rrc[n]$) and the sinusoidal Hilbert-pair basis ($\{\cos, \sin\}$):

$$g_s[n] = f_r[-n] = \begin{cases} rrc[n]\cos(2\pi f_c n) & \text{for } s = r = 1 \\ rrc[n]\sin(2\pi f_c n) & \text{for } s = r = 2 \end{cases}, \tag{3.18}$$

where fc is the pass-band's center frequency, and $rrc[n]$ is given by [130]



(a) $g_1[n]$ (b) $g_2[n]$

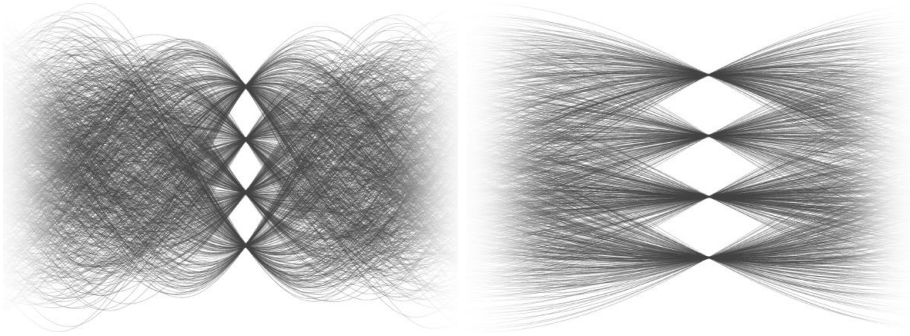


(c) $|G_1(f)|^2$

Figure 3.5: On the top, orthogonal filter pair in conventional 2D-CAP for $f_c = F_B$ and $\beta = 0.1$. On the bottom, $G_1(f)$ magnitude squared. See 3.18.

$$rrc[n] = \begin{cases} \frac{1}{\sqrt{T}}\left(1-\beta+4\frac{\beta}{\pi}\right) & \text{for } n=0 \\ \frac{\beta}{\sqrt{2T}}\left(\left(1+\frac{2}{\pi}\right)\sin\left(\frac{\pi}{4\beta}\right)+\left(1-\frac{2}{\pi}\right)\cos\left(\frac{\pi}{4\beta}\right)\right) & \text{for } n=\pm\frac{T}{2\beta} \\ \frac{1}{\sqrt{T}}\frac{\sin\left(\pi\frac{n}{T}(1-\beta)\right)+4\beta\frac{n}{T}\cos\left(\pi\frac{n}{T}(1-\beta)\right)}{\pi\frac{n}{T}\left(1-\left(4\beta\frac{n}{T}\right)^2\right)} & \text{otherwise} \end{cases}, \tag{3.19}$$

Figure 3.5 shows an example of orthogonal pulses, $g_s[n]$, as described in 3.18. Now, expressing the transmitter output, y, according to 3.18:



(a) CAP (b) Raised cosine

Figure 3.6: Recovered eye-diagrams over one symbol-period of 4-level conventional 2D-CAP for $f_c = F_B$ (a), and for raised cosine pulse-shaping (b); both with $\beta = 0.1$.

$$\begin{aligned} y[n] = &\sum_{k=-\infty}^{\infty} A_1[k] rrc[n-kT] \cos(2\pi f_c(n-kT)) + \\ &+ \sum_{k=-\infty}^{\infty} A_2[k] rrc[n-kT] \sin(2\pi f_c(n-kT)) \end{aligned} \tag{3.20}$$

The above expression proves that CAP behaves as a regular quadrature modulator (with the exception of a $2\pi f_c kT$ phase shift, which turns irrelevant from the $rrc[n]$ function perspective, as well as for maintaining the orthogonality condition) without requiring analog mixers and with accurately adjustable center frequency and SE through β. Notice that the use of raised-cosine is not strictly necessary for attaining zero ISI as long as fc is an integer multiple of $F_B = \frac{1}{T}$.

Eye pattern : as a linear system that CAP is, there is not TX/RX mixing element realizing up-/down-conversion of the data signal; but rather the pulse-shaping filter itself defines a pass-band impulse response. Consequence of this, the output signal after matched filtering at the receiver remains in pass-band, where the demodulation takes place, and hence displaying eye diagrams with unusually narrow horizontal opening (see comparison in Figure 3.6).

Highly accurate timing recovery and sampling point estimation are required if conventional slicing is used for signal demodulation, motivating

the study of alternative methods for CAP filter design that manage to relax such stringency [141]. Nevertheless, considerably more robust demodulation is attained if statistical tools are contemplated for level identification. The reason for this performance improvement is subtly explained by 3.20 for conventional 2D-CAP, making visible that the horizontal eye closure is dictated by the $2\pi f_c kT$ phase shift resulting from the convolution. Deviating from the maximum variance sampling instants thus translates into cross-dimensional coupling, which shows as a standard rotation of the complex 2D constellation. Simple clustering algorithms based on Euclidean distance such as *k-means* are able to double the sampling-deviation tolerance without requiring special filter design [96].

Peak-to-mean envelope power ratio : similarly to other multiplexing techniques, a linear superposition of independent waveforms takes place in ND-CAP for $N > 1$, leading to local interference instants with abnormally high magnitudes. In the electrical domain, these power peaks make the signal susceptible to non-linear effects in the power amplifiers, and exacerbate the quantization noise when the full dynamic range of the signal is accommodated within the DAC/analogue-to-digital converter (ADC) resolution. In addition, because fiber nonlinearities are also dependent on the instantaneous field magnitude, large values of PAPR reduce the nonlinear average-power threshold, then forcing to operate the system at lower OSNRs. Accordingly, we next elaborate on the PAPR performance of ND-CAP signals as an essential step in the design of any multiplexed transmission system. Without loss of generality, we employ the peak-to-mean envelop power ratio (PMEPR) parameter, which is the base-band equivalent of PAPR[12] [142, 143]. In all CAP cases, the filters are 14-symbol long and the statistical analysis is performed over 20000 samples of 300 symbols each.

Figure 3.7 shows the complementary cumulative distribution function (CCDF) of the PMEPR for (i) a set of realizations of 1-to-5D-CAP systems, all of them calculated through the previously mentioned *minimax* algorithm constrained to 3.12 with $\beta = 0.85$; and for (ii) uniformly bit-loaded 256-carrier discrete multitone (DMT)[13], well-known for its extraor-

[12]Consequence of the additional sinusoidal term after up-conversion to the carrier frequency, the PAPR is twice as high as the PMEPR. A ±3-dB correction term can be applied for back-forth conversion according to $\text{PMEPR}(x[n])|_{dB} = \text{PAPR}(x[n])|_{dB} - 3dB$.

[13]DMT technique is a modified version of OFDM in which Hermitian symmetry is forced to the input of the *ifft* block typically employed as the filter-bank for signal generation. As a result, the transmitter output is purely real, hence allowing for exploiting

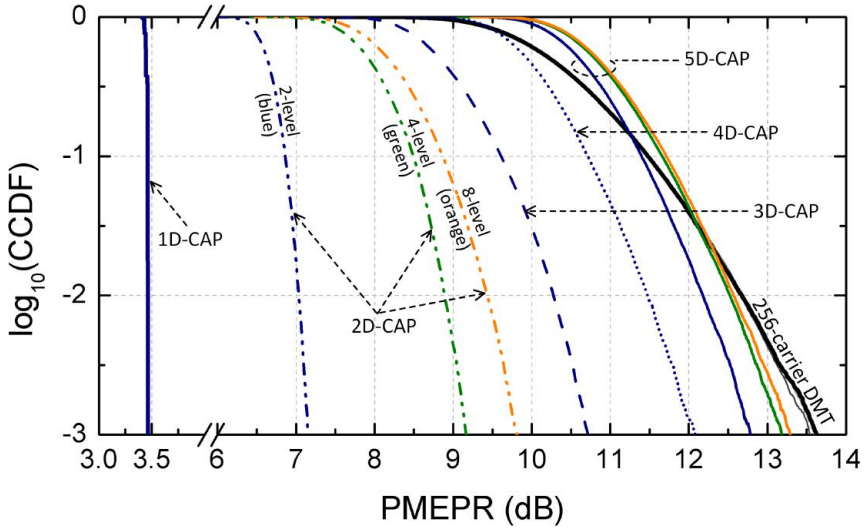


Figure 3.7: CCDF of PMEPR for concrete realization of 1-to-5D-CAP systems with 2- (orange), 4- (green), and 8-level (blue) amplitude-modulation per dimension, and for uniformly bit-loaded DMT.

dinary PMEPR [143] and serving as worst-case reference. In all scenarios, 2-, 4-, and 8-level amplitude-modulation per dimension are simulated, although some curves are intentionally removed due to the little additional information provided, and for the ease of visualization.

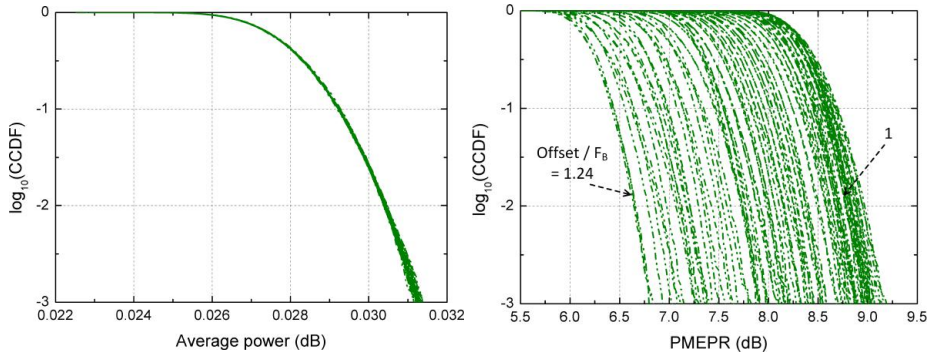
A quick comparison with the 2-level 1D reference makes clear the negative influence of $N > 1$ CAP at all CCDF levels. Owed to the high sensitivity of the PMEPR to the modulation order in single/few-carrier systems, this difference substantially increases for 4- and 8-level modulation, as it is visible in 2D-CAP with up to ≈3-dB additional PMEPR in the latter case at CCDF=10^{-3}. Predictably, the relative and absolute magnitude of this penalty experiences saturation due to two phenomena: the stabilization of the average power of the peak-normalized signal for incremental constellation size, which governs the differential penalty; and the fact that the PMEPR is predominantly dominated by the incoherent waveform superposition for high values of N, making the penalty increasingly insensitive to the modulation level (see how the performance for all orders converge in 5D-CAP). An extreme case of the latter observation is DMT, whose PMEPR performance is primarily determined by the amount

the potential of OFDM in scenarios employing intensity modulation.

of superimposed subcarriers (typically in the order of hundreds), appearing totally invariant with the modulation per carrier. Nevertheless, the PMEPR clearly raises with N despite an apparent drop in the incremental step; and consequently, 4/8-level 4D-CAP (not visible) and the three cases in 5D-CAP follow DMT's curve with notable proximity; evincing the need to pay serious attention to PMEPR-driven performance degradation even in single-carrier configuration when high-order modulation and multiple dimensions are to be employed.

Furthermore, the CCDF profiles of CAP and DMT are noticeably disparate, where the former decreases more steeply. The consequence of this observation turns of importance in *clipping*, a commonly used technique to alleviate PMEPR-driven penalties by which all values surpassing a predefined threshold are clipped out [144]; and hence a proportional amount of information is lost. In this regard, flatter curves are more prone to clipping, as the dynamic range of the signal can be reduced with hardly any information loss and *vice versa*; demonstrating that *clipping* in ND-CAP signals is not as effective as in DMT/OFDM, albeit the trend slightly equalizes with increasing N. Nevertheless, CAP exhibits lower average PMEPR for all studied dimensions and any modulation order, being $N = 4$ the largest dimensionality experimentally demonstrated in optical communications as yet [145]; and making CAP's (or PAM's) receiver sensitivity provenly superior in conditions where nonlinearities are fundamental limiting factors [146, 147]. However, the picture changes substantially when band-limited or frequency-selective channels are taken into account, where spectral flexibility is most-wanted property instead [90, 96].

Finally, we observed that the PMEPR is heavily dependent on the band's center frequency; potentially changing by more than 2 dB between certain implementations. As illustrative example, Figure 3.8 shows the PMEPR performance of 4-level original 2D-CAP for 200 center frequencies (f_c in 3.18) ranging from 1 to 2 times the symbol rate F_B (right); and the CCDF of the respective average power (left), evincing that it is exclusively the peak power what drives the extra degradation. The filters are 25-symbol long and the statistical analysis is performed over 20000 samples of 500 symbols each per frequency offset case. We observed that the offset-driven PMEPR penalty is periodic every F_B with mirrored symmetric every $0.5{\times}F_B$ at exactly F_B periods, thus evaluating 1-to-$0.5{\times}F_B$ would anyways cover the possible dynamic range entirely. For example, the curve at $1.1{\times}F_B$ precisely coincide with the ones at 0.9, 1.9 or $2.1{\times}F_B$. Even though locking the band to the best-performing frequency range is



(a) (b)

Figure 3.8: In b, CCDF of PMEPR of 4-level original 2D-CAP system for 200 different center frequencies ranging from 1 to 2 times the symbol rate F_B; in a, the CCDF of the average power.

definitely doable, it should be noted that CAP bands are often applied a β-dependent offset to place them near DC with the purpose of maximizing the electrical bandwidth usage and avoid over-filtering effects to the largest extent. In these cases, and those where bands may exhibit arbitrary offsets like in next-described MultiCAP systems, this impairment factor may be worth consideration.

3.2.1.2 MultiCAP

One relevant conclusion from the previous section is that CAP bands are not restricted to a particular center frequency, but rather to a minimum spectral occupancy determined by 3.13. That is, on condition that 3.14 is satisfied, several independent multi-dimensional bands can be arbitrarily arranged across the available bandwidth. This is the core idea in MultiCAP, where the electrical spectrum is subdivided into a user-defined number of low-baudrate bands, enabling highly flexible spectral control and hence proven robustness against frequency-selective distortions (e.g. CD [96]) or in strong band-limited channels (e.g. visible light communications [148], or cost-effective fiber-optic short-haul systems [133,135]) without jeopardizing CAP's fundamental signal properties. Figure 3.9 illustrates a basic example of channel-dependent power-loading in a 3-band 2D-MultiCAP with same bandwidth per band.

Mathematically, each of the bands in MultiCAP accepts identical end-

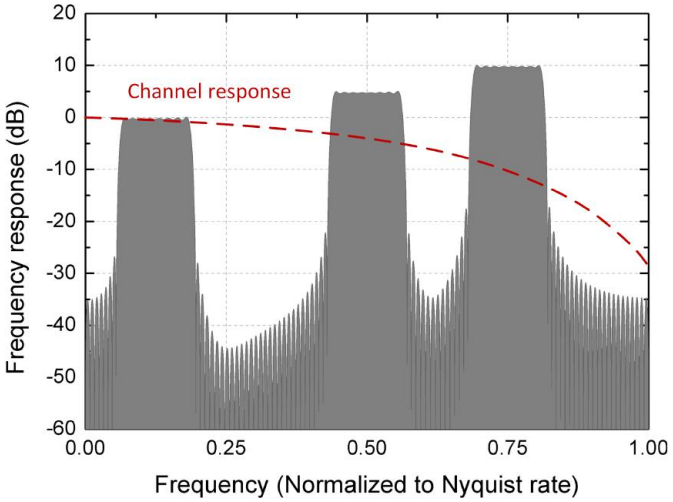


Figure 3.9: Frequency response of one realization of 3-band 2D-MultiCAP. The spectral occupancy for all the bands is 12.5% of the Nyquist rate, and they are centered at 12.5%, 50% and 75% of the Nyquist rate. Their power gain is adjusted to compensate for the loss induced by an emulated channel response (red dashed line) at the respective center frequencies.

to-end formulation as standard CAP (see 3.3), and consequently, the full system can be viewed as an M-fold concatenation of independent CAP instances. This drives the attention to the major implementation challenge in MultiCAP system with large number of bands and dimensions, the processing complexity; directly related to the required number of filters and their taps. To quantify the bare minimum number of coefficients per filter, we start by adapting the sampling theorem in 3.14 to multiple bands:

$$\begin{aligned} F_{S;min} = 2BW_{min} =& 2\sum_{b=1}^{M}\left(\frac{F_{B;b}}{2}N_b(1+\beta_b)+\gamma_b\frac{F_{B;b}}{2}\right) = \\ =& \sum_{b=1}^{M} F_{B;b}\left(N_b(1+\beta_b)+\gamma_b\right) \end{aligned} \tag{3.21}$$

where M is the total number of bands, $F_{B;b}$ is the symbol rate of the band b with $b \in \mathbb{N}$, β_b is the respective roll-off factor, N_b the number of dimensions, and γ_b is the fraction of $\frac{F_{B;b}}{2}$ standing for the frequency guard-band on the lower frequency side of the corresponding band (custom frequency offsets shall be included in this parameter). Now dividing 3.21 by the symbol rate

per band $F_{B;b}$ and rounding to the immediate superior integer, we obtain the associated number of samples per symbol:

$$S_{sym;b} = \left\lceil \frac{\sum_{b'=1}^{M} F_{B;b'} \left(N_{b'}(1+\beta_{b'}) + \gamma_{b'}\right)}{F_{B;b}} \right\rceil \tag{3.22}$$

which in the case where all parameters are identical for all bands, 3.22 simplifies to:

$$S_{sym;b} = \lceil M \left(N(1+\beta) + \gamma\right) \rceil \tag{3.23}$$

The influence of sub-band partitioning ($M > 1$) and possible guard intervals ($\gamma > 0$) on the number of samples per symbol is clearly revealed in 3.22, becoming particularly explicit for the scenario described by 3.23. Considering 3.23 for simplicity, we observe that even when maximum spectral compactness is assumed for $S_{sym;b}$ minimization ($\gamma = \beta = 0$), the number of samples per symbol is increased M-fold per additional band. Therefore, reworking the calculations in Section 3.2.1.1, M-band ND-CAP uses MN times longer pulse-shaping filters than 1-band 1D-CAP for achieving the same SE (MN S/symbol); resulting in a total of $(MN)^2$ times more filter coefficients to be processed for signal de/multiplexing. For example in a 6-band 2D-CAP configuration, assuming 15-symbol long filters so that the BER error floor keeps $\leq 10^{-9}$ for $\beta = 0.1$ in all bands [96], the total number of extra taps disregarding equalization equals $(6{\cdot}2)^2{\cdot}15 = 2160$; approximately 7-to-8 times the amount required for bulk frequency-domain CD compensation of 1000-km SSMF transmission [149].

Besides the additional complexity, two further pertinent conclusions are drawn from the above equations concerning digital signal generation. First, the over-sampling factor per band, M, is not strictly needed for its perfect generation, but rather a consequence of the required Nyquist sampling rate for the ensemble/multi-band spectrum. Thereby the number of samples per symbol and hence the spectral occupancy is variable and potentially different among bands. Assuming a band-limited channel for which BW_{min} is wished invariant, the eligible range of symbol rates is directly proportional to M; providing the system with an additional tool (not present in OFDM/DMT) for rate adaptation and spectral shaping with M-driven granularity. The quantitative influence of M on the per-band spectral occupancy is made explicit through the symbol rate by imposing $F_{B;b}S_{sym;b} = 2BW_{min}$ to be constant and resolving for 3.22.

Second, the minimum sampling frequency satisfying 3.22 given a fixed BW_{min} converges to the Nyquist rate ($2BW_{min}$) as MN increases. This is

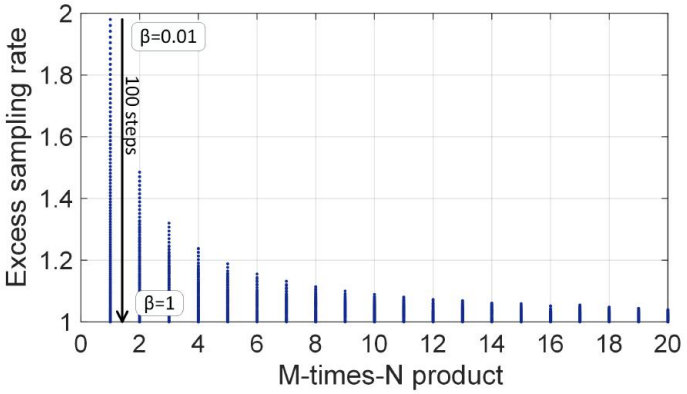


Figure 3.10: Excess sampling rate versus MN product for the system configuration described by 3.23 with $\gamma = 0$. Hundred values of β ranging from 0.01 (highest values) to 1 (flat bottom line at 1 excess sampling rate) are evaluated.

because the deviation from exact number of samples per symbol induced by the ceiling function in 3.22 becomes progressively negligible as the operand grows. The excess sampling rate is then given as the ratio between $S_{sym;b}$ with and without ceiling function; which is illustrated in Figure 3.10 for the simplified configuration in 3.23 with $\gamma = 0$ (maximum SE). For example, we observe that conventional 1-band 2D-CAP with $\beta = 0.01$ requires 1.5 times higher sampling rate than theoretically necessary; however, if the number of bands is increased to 3 while maintaining the bi-dimensionality ($MN = 6$), the excess rate goes down to 1.15. This means that MultiCAP allows for signal generation across the full-bandwidth of the DAC with more relaxed sampling clocks; or from a different perspective, more efficient use of the available bandwidth is attainable.

Peak-to-mean envelope power ratio : exactly like multiplexing subcarriers or increasing the CAP's dimensions do, assembling independent sub-bands increases the PMEPR. Whereas this is a readily foreseeable point, the complete characterization of all the possible combination of parameters (M, N_b, β_b, γ_b, amplitude modulation order, and symbol rate) is an intractable endeavor; hence some pragmatic selectivity is mandatory. In this regard, we have demonstrated the quadratic growth of the processing complexity with M and N_b in the optimum-SE scenario (see 3.23 and

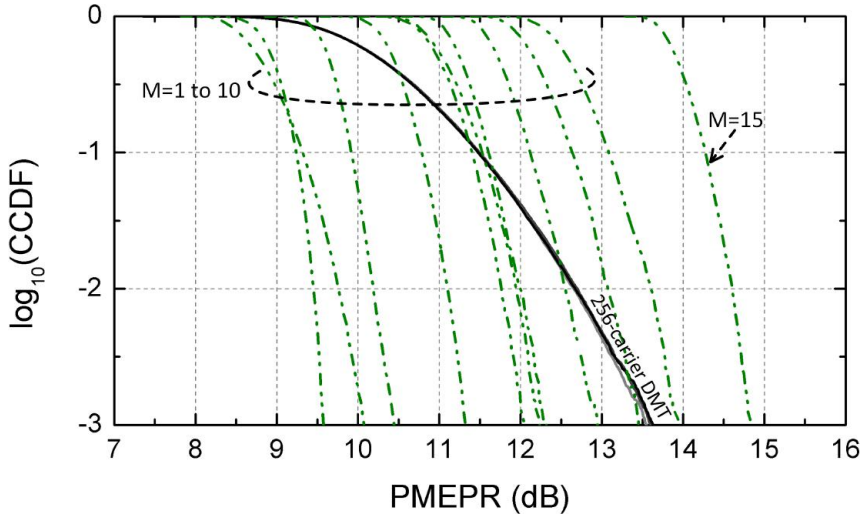


Figure 3.11: CCDF of PMEPR for 1-to-10-band (plus 15-band case) 4-level original 2D-CAP system. Uniformly bit-loaded DMT is included for reference.

subsequent elaboration), from which N_b has been demonstrated to raise the per-band PMEPR to equivalent ranges as DMT beyond three dimensions (see Figure 3.7). On the other hand, because the primary purpose of MultiCAP is providing spectral flexibility, the number of bands is wished $M \geq 2$ where the upperbound is determined by the desired granularity and the associated processing complexity. Therefore, because of these observations, we reduce the analysis to zero guard-band ($\gamma = 0$) for maximizing the SE, a practical $\beta = 0.1$ for which 10 symbols per filter suffice for BER error floor at 10^{-9}, $N = 2$ in order to have enough PMEPR budget to accommodate $M > 2$ incurring neither excessive degradation nor complexity, and up to $M = 10$ for allowing independent spectral shaping in fractions of 10% of the total bandwidth. For simplicity, $F_{B;b}$ is considered the same in all bands (≈3% of the sampling rate), bit- and power-loading are uniform distributed across all bands, and conventional 2D-CAP is employed for signal generation.

Referencing single-band CAP, ≈4.5 dB PMEPR penalty is observed at CCDF=10^{-3} when the number of bands is increased to 10, with an absolute maximum of ≈14 dB. Besides the expected behavior, three other observations are worth mentioning. First, the curves' profile are independent of M (except for the outlier at M=2) and remarkably steep (≈2 orders of mag-

nitude in CCDF per dB in PMEPR); underscoring the little efficiency of *clipping* for 2D configuration in both CAP (see Figure 3.7) and MultiCAP (the tendency appears mainly dependent on N). Second, the incremental penalty per additional M is variant, proving the existence of an optimum ΔPMEPR-to-ΔM ratio. For example 5, 6, and 7 bands perform in less than 0.25 dB from each other at CCDF=10^{-3} and up to $\approx$1.5-dB better than 256-carrier DMT; strongly encouraging $M \in \{5, 6, 7\}$ as long as the corresponding spectral partitioning granularity (from 14.2% to 20% of the total bandwidth) and the processing complexity meet the target requirements. Third, unlike in DMT/OFDM, the degree of spectral flexibility in MultiCAP systems is traded on PMEPR and complexity without jeopardizing the SE. This enables efficient system designs better adjusted to the needs in spectral control. For instance, if 25% partitioning granularity was hypothetically enough in a given scenario, we could use the above-described configuration with just M=4; achieving similar SE than 256-carrier DMT with more than 2 dB better PMEPR performance at CCDF=10^{-3}. Further performance gains could be achieved by optimally adapting the bandwidth per band to the end-to-end response (e.g. finer granularity to combat higher attenuation selectivity).

3.2.1.3 C-MultiCAP

Elementally based on MultiCAP system, C-MultiCAP enables its use in coherent communications, aiming for longer transmission distances and higher throughput. Besides the LO-driven sensitivity improvement, quadrature and polarization components[14] are now available to carry M-band ND-CAP signals; providing 4-fold extra dimensions whereby the SE is multiplied, and multi-dimensional modulation formats of higher order are made possible.

Irrespective of the application, MultiCAP signal represents the building block in C-MultiCAP generation, and so the same mathematical formulation, signal properties and implementation constraints per coherent-dimension hold (see Section 3.2.1.2). Differently however, C-MultiCAP transceivers are devised to compensate for a handful of essential distortions typically associated with polarization-multiplexed coherent communication: carrier frequency offset plus phase noise, and quadrature plus polarization mixing. Whereas electronic mitigation of such impairments

[14]To avoid confusion with CAP dimensions (N), quadrature and polarization components (4 in total) are referred to as *coherent-dimensions*.

is already a mature subject in both single-carrier configurations [85, 150] and OFDM systems [151], C-MultiCAP presents original characteristics that impede the straight application of existing DSP structures. Adapting those individual algorithms and structures according to the signal properties is thus obligatory, and as it is often the case, it implies taking specific actions in the transmitter side. Next, we elaborate on the transmitter-side modifications and generation constraints adopted in C-MultiCAP to enable impairment compensation and successful demodulation.

Frequency offset and phase noise

Commonly operated in intradyne configuration, coherent detection results in a time-varying phase distortion that couples both quadrature components in a random and dynamic manner. Typically in single-carrier approaches employing AMF, this residual phase variation is compensated via constellation-driven second-order DPLL [85]; whereas in OFDM, special training symbols plus pilot carriers are transmitted for carrier frequency and common phase noise estimation respectively [151]. Notwithstanding, none of these approaches suit the functional requirements of MultiCAP-based signals for the following reasons:

- High-SE systems featuring spectral fragmentation rely upon exceptionally selective frequency demultiplexing. Particularly in C-MultiCAP, band separation is performed through a set of non-overlapping matched filters per coherent-dimension, which are in turn necessary for sustaining the N_bth order dimensionality and ISI-free communication (see Section 3.2.1.1 and Section 3.2.1.2). That is, RX CAP filtering is not just a digital means of isolating the sub-bands, but also indispensable front computation to enable any subsequent processing towards successful demodulation. Failing to maintain exact TX-RX filter matching in CAP-based systems along with quadrature coupling, causes the combined effect of: (i) intra- and cross-coherent-dimension inter-band interference (IBI), (ii) asymmetric band attenuation, and (iii) incorrect cancellation of the ISI inherent to CAP signals. As a result, in the absence of prior compensation of carrier's frequency offset and phase noise, the standard real-valued CAP filters work as $4\times2\times N_b\times$time-dimensional[15] coupling functions;

[15]The dimensionality is given by: 4 coherent-dimensions (polarization-mixing is assumed), 2-bands per coherent dimension (worst-case cross-dimensional IBI), N_b CAP

compelling to process the entire signal spectrum in an intractable blind adaptive equalizer. This represents the major difference with single-carrier digital coherent receivers, where filtering and equalization is first step to boost the SNR and thereby improving the estimation accuracy of the residual phase.

- Owed to reasons relating to electro-optical generation stability and digital processing ease plus performance, training symbols are designed so that the occupied bandwidth and the average power remain as constant as possible. In this regard, because OFDM/DMT modulation is performed directly in the frequency domain, time-sensitive training symbols with different properties for frequency recovery [152] can be readily inserted and, at the same time, jointly satisfying method-specific and scenario-dependent spectral and power constraints. For example, odd sub-carrier suppression plus power loading to create even symmetry [153] while compensating for the power loss associated to the nulling process and adapt to the channel response. Moreover, the PMEPR in OFDM/DMT systems is insensitive to the modulation order as long as the number of carriers is in the order of hundreds; in turn making feasible the use of very low-bandwidth dummy sub-carriers spread across the spectrum for tracking the common phase noise with affordable overhead [151].

  Modulation in MultiCAP, on contrary, is carried out in time-domain on a band-by-band basis, whose center frequency is to remain adjustable, and whose estimated minimum spectral occupancy is bounded by processing complexity and PMEPR between 14.2% and 20% of the electrical modulation bandwidth (see Section 3.2.1.1). Clearly on this account, the design of effective training symbols for frequency offset compensation meeting such requisites in spectral flexibility is utmost challenging, and the use of dedicated sub-bands for tracking the phase noise is prohibitive from the overhead standpoint.

In summary, coherent DP-MultiCAP transmissions need blind estimation of the residual frequency offset and phase noise at the beginning of the RX DSP chain, thereby increasing the isolation ratio among intra-band CAP dimensions; and minimizing the random/noisy contributions from intra- and cross-dimensional IBI. As it will be reasoned in the next point, these benefits facilitate per-band equalization with standard single-carrier

dimensions per band, and time (ISI).

tools when the characteristic parameters of the constituent MultiCAP signals are identical.

For this matter, we propose inserting one or multiple unmodulated frequency pilot tones (FPTs) as low-overhead gauge for unambiguous, continuous and blind phase tracking; a simple method whose effectiveness and robustness as compared to other alternatives has been demonstrated in several scenarios [46, 122, 154–158]. The reader is referred to Appendix C for detailed description of the FPT technique and general implementation concerns.

Quadrature and polarization mixing

Even after ideal frequency offset and phase noise compensation, it exists certain degree of cross-dimensional coupling originating from polarization mixing plus PMD, residual CD (prior bulk CD compensation in understood), and possible filtering effects specially on the outer sub-bands. Compensating for those effects demands the transmitted signal to present certain deterministic relations among coupled dimensions, which can later be exploited at the receiver as reference to recover all independent signals. Clear examples are the stationary characteristics of the constellation diagrams employed in popular blind equalization techniques like CMA or RDE [85], space-time codes for polarization demultiplexing [159], or the use of 2D training sequences jointly carried by both quadrature components (e.g. n-ary QAM or CAZAC) [160, 161]. For this reason, the parameters of the constituent MultiCAP signals must be neither random nor mutually independent, but rather describe a signal structure which facilitates effective equalization with as little complexity as possible.

Accordingly, we constrain MultiCAP's parameters (M, N_b, β_b, γ_b, amplitude modulation order, and symbol rate) to be variable but homogeneous across all bands, a decision resulting from various other assumptions and requirements: (i) end-to-end filtering impairments are similar in all coherent dimensions, so that the same spectral arrangement equally applies in an optimum manner; (ii) all the operative frequency ranges (not disabled by filtering restrictions) within the available electro-optical bandwidth are to be used for SE maximization; and (iii) lessening the maximum absolute number of overlapping bands is desirable to cut down the plurality of interfering dimensions, and thus to simplify equalization. In addition, we impose CAP filters' impulse responses to be identical in overlapping bands which, leveraging the structural constraints, conveniently allows for building N_b

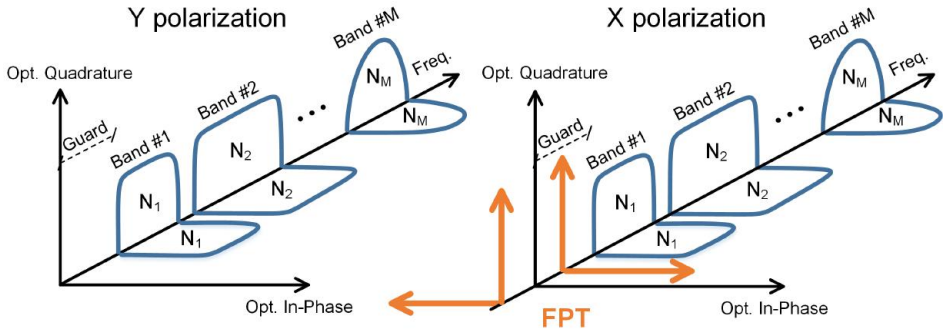


Figure 3.12: Conceptual illustration of C-MultiCAP signal spectrum as devised in this investigation. Both polarization and quadrature components are represented.

equivalent DP complex base-band signals per band. The use of well-known algorithms in standard single-carrier digital coherent receivers [85, 150] is thereby enabled per CAP dimension (see Section 3.2.2), making possible the scalability of N_b at a proportional increase in processing load.

Final C-MultiCAP transmission system : the spectrum of C-MultiCAP signal is conceptually depicted in Figure 3.12, incorporating the above-mentioned structural requirements. All coherent-dimensions carry identical MultiCAP structures, where M, N_b, β_b, γ_b, amplitude modulation order, and symbol rate are allowed to vary within margins determined by a compromise between spectral flexibility, rate adaptation granularity, processing complexity and PMEPR performance. It is important to remark that the structural restriction does not prevent transmitting independent data per dimension.

One single-sideband radio-frequency FPT is digitally inserted (see Appendix C) between DC and the first spectral component of band #1 in one polarization. A frequency guard-band of $\leq$500-MHz is allocated for its insertion[16], representing the only efficiency loss per coherent-dimension as compared to standard MultiCAP; less than 3% overhead considering $\geq$20-GHz modulation bandwidth. The reasons for this layout are:

- Near DC (carrier frequency after up-converion) the accumulated pass-band and low-pass filtering distortions are decreased, thus improving

[16]The concrete value depends on the phase noise characteristics and the filter employed for FPT isolation.

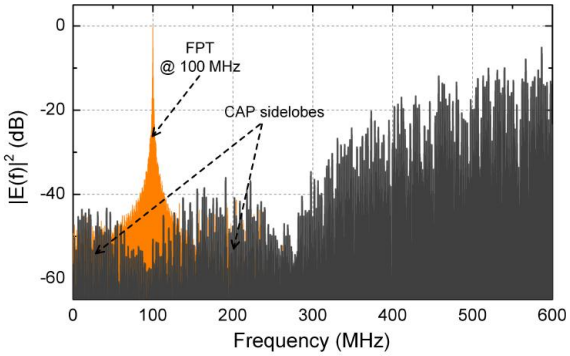


Figure 3.13: Simulated spectrum of 2D-CAP signal with ≈300-MHz guard-band with respect to DC (grey); and one 100-MHz FPT (orange), right between sidelobes for interference minimization.

the estimation performance over longer transmission distances. Additionally, because DC is the symmetry center of the signal spectrum, the maximum distance to the highest modulated frequency is minimum, which reduces the gravity of the phase misestimation caused by group-delay mechanisms such as CD.

- Strong phase noise and broad filters increase the sidelobes-FPT interference, hence it is convenient to allocate ample guard-band range (up to 500 MHz in our case) to shift in the lower-magnitude sidelobes of band #1 whenever necessary. However, because the sidelobes' attributes vary with different MultiCAP parameters, the FPT may undergo severe interference even for high offsets. To solve this, we digitally aligning the FPT center frequency to the spectral zero closest to DC, thus minimizing the cross-talk, stabilizing the phase estimation reliability, and allowing for better pilot-to-signal ratios (PSRs). Figure 3.13 shows a clarifying example.

- Because the FPT period is hundreds of times longer than the expected differential group delay (DGD) in MAN networks (see footnote 3), the equivalent polarization-coupling matrix acting on the FPT can be approximated by a real-valued rotator. By transmitting only one FPT in one polarization, coarse π-ambiguous estimation of the mixing factor can be performed through blind search upon the criterion of maximum constructive interference (see Section 3.2.2). In addition to the prior compensation of the phase noise and the residual frequency

offset, this operation potentially improves the performance and the tap-convergence speed of the adaptive equalizer.

Considering the guard-band for FPT insertion, the maximum gross throughput is thus ≈97% of 4×MultiCAP, exactly the same 4-fold factor by which multiplexing complexity is multiplied. The extra computational cost of FPT insertion should also be added, since it represents a core functional element in C-MultiCAP. In terms of PMEPR, and benchmarking MultiCAP, overall reduction is experienced consequence of the addition of the orthogonal coherent-dimensions. The relative performance against DMT is thus expected to be maintained within the same range (see Figure 3.11); although their disparate PMEPR statistical properties, and the influence of the amplitude modulation level in C-MultiCAP suggest scenario-specific analysis. In this regard, although the influence of the FPT itself on the PMEPR is negligible with typical PSR around -15 dB, the particular frequency offset induced on the MultiCAP signal for allocating the FPT guard-band may be determinant and shall be contemplated.

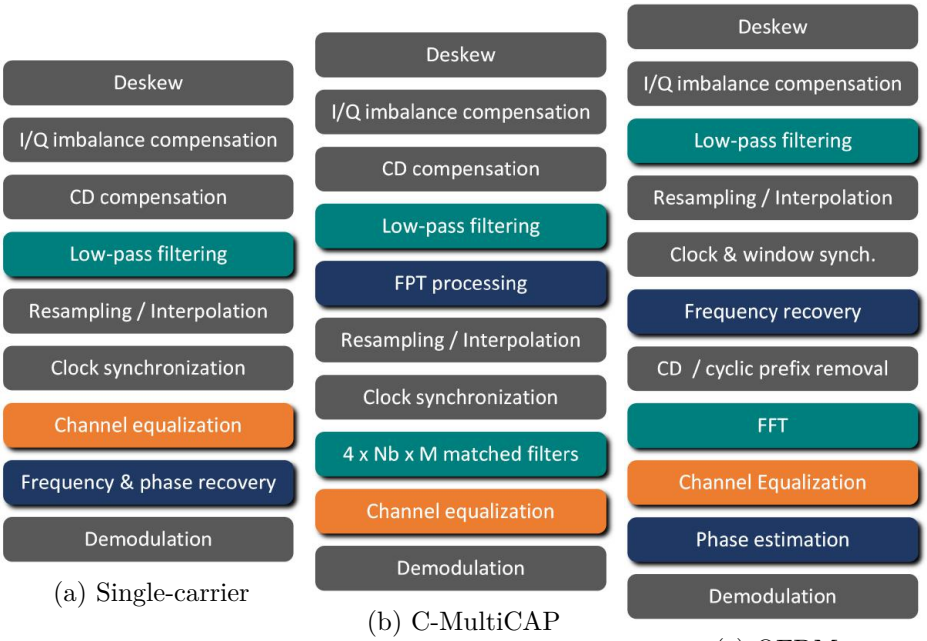
3.2.2 Detection

Enabled by FPT technique and structural restrictions, multi-dimensional and spectrally flexible C-MultiCAP signals can solely leverage DSP for their successful transmission in coherent communications. Whereas the algorithms in C-MultiCAP and standard single-carrier systems are very much alike, their succession is disposed similarly to OFDM's in the former; resulting in a mixed organization worth detailing. In the following subsection, the arrangement of RX DSP algorithms for C-MultiCAP detection is discussed as employed in this investigation; from which the FPT-based phase tracking method and the multi-dimensional adaptive equalizer are further developed as the primary modifications with respect to standard single-carrier's DSP.

3.2.2.1 DSP structure

Figure 3.14 shows the schematic diagram of the RX coherent DSP for single-carrier (a), C-MultiCAP (b), and OFDM (c) systems. It is visible that C-MultiCAP processing represents an intermediate compromise with strong resemblance with both other alternatives.

The first four blocks in C-MultiCAP are similar to single-carrier's, that is, after initial front-end corrections, bulk CD compensation is applied to re-



(a) Single-carrier (b) C-MultiCAP (c) OFDM

Figure 3.14: Digital coherent receiver structures for single-carrier, C-MultiCAP and OFDM signals. Color-matched blocks across structures perform operations of similar purpose.

lieve the complexity of the posterior equalizer, followed by broad-band low-pass filtering that removes out-of-band noise and hence improves the overall SNR. At this stage, the residual frequency offset from intradyne detection and the phase noise is estimated by means of FPT. Conditional upon correct timing recovery[17], FPT processing mitigates the inter- and intra-coherent-dimension IBI potentially occurring in the successive $4{\times}M{\times}N_b$ matched filtering processes, and permits the correct separation of the N_b dimensions per band. In addition, when the FPT is inserted in one polarization exclusively, coarse estimation of the polarization mixing angle is possible via blind search; with the cost function determined by the peak magnitude of the field's Fourier transform on the employed polarization,

[17]Because clock recovery is needed before having access to the modulated waveforms, frequency-domain approaches are mostly adequate. In this concern, techniques exploiting FPTs have been long proved successful [162] also in similar multi-carrier scenarios [163], although their utilization in C-MultiCAP systems may require adaptation given the variability of the bands' allocation.

and maximum absolute as break condition. See Subsection 3.2.2.2. Coarse polarization de-rotation speeds-up the convergence time of the following adaptive equalizer.

Afterwards, relying on the the fact that all 4 coherent-dimensions employ the same MultiCAP parameters and filters per band (see the explanation after Figure 3.12), every CAP dimension in every band is grouped up with the corresponding one from the other coherent-dimensions. The aggregate entity is treated as an equivalent DP complex base-band signal, and then transferred to the equalization block ($M \times N_b$ in total) for band-by-band linear impairment mitigation. Three alternative approaches may be employed: (i) the algorithm is run on only one equivalent DP complex base-band and then applied to the remaining dimensions, then reducing the processing complexity; (ii) the channel is individually calculated for each dimension, improving the performance and robustness; and (iii) the estimated impulse responses of $\geq$2 CAP dimensions are averaged and the outcome is applied to the entire band, which diminishes the influence of the noise in low SNR conditions. Irrespective of the option, because the input waveforms to the equalizer have been previously compensated for frequency offset and phase noise, the use of decision-directed error functions is possible; which can further reduce the convergence time [164]. See Subsection 3.2.2.3.

3.2.2.2 FPT processing

Depending on the number of pilots and their distribution, the processing steps for phase estimation and the associated performance will change, as well as their availability for additional applications like clock recovery or coarse polarization decoupling. For concision and generality, in this part we elaborate on the DSP blocks and applications particularizing to one single-sideband radio-frequency FPT digitally inserted into one of the orthogonal polarization states. This arrangement is employed in the experimental demonstration discussed Section 3.3.

Coarse polarization decoupling : Along propagation over optical fiber, the polarization state of the field changes. When this field impinges upon the static front-end polarization splitter, inescapable signal mixing occurs, which for the case of an FPT with period$\gg$DGD can be well modeled with a 2$\times$2 real-valued rotation matrix. This being the case, if the FPT is inserted only in one polarization, the mixing angle can be estimated by searching over a set of test rotators the one that yields maximum

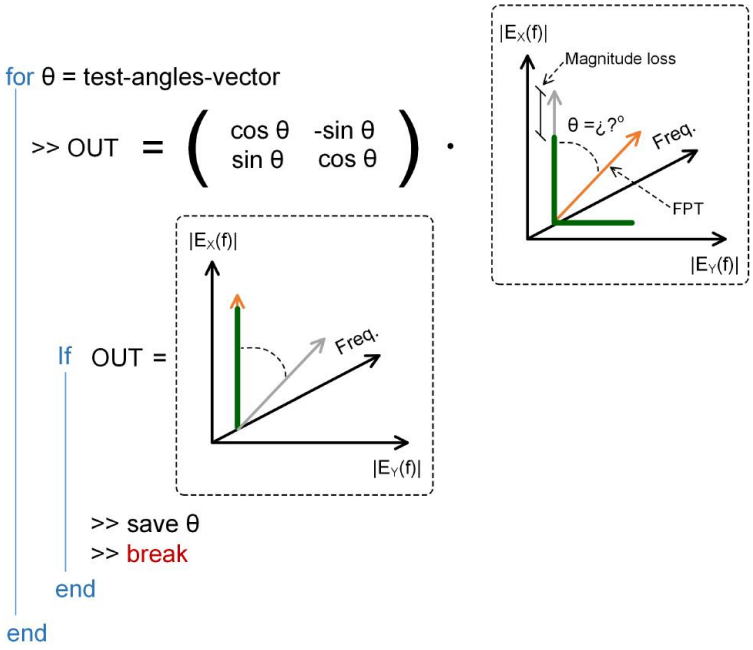
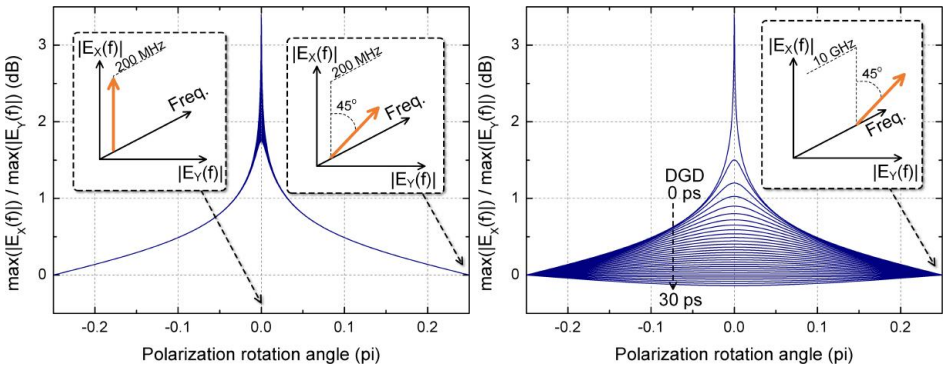


Figure 3.15: Graphical representation of the iterative algorithm for coarse estimation of the signal's polarization mixing angle. Top-right inset figure conceptually illustrates the power-leakeage effect of an arbitrary low-frequency FPT on the spectra of both orthogonal polarization fields when the mixing is not corrected for. Likewise, the middle inset figure represents the target power coupling.

peak magnitude on the target polarization. The inherent π ambiguity can be compensated after or during equalization by means of a correlated DP training symbol. Figure 3.15 shows a clarifying illustration of how the algorithm proceeds, the effect of mixing on the magnitude spectra of both polarization, and the target coupling ratio. Note that for typical PSRs, the tone's peak prominently rises over the maximum signal level perfect polarization alignment (e.g. Figure 3.13), hence the maximum-peak criterion holds without requiring specific FPT isolation.

Because this technique relies on keeping a very high ratio between the FPT period and the DGD, it is important to assure compliance given the expected scenario-specific PMD values (see footnote 3) and the concrete FPT allocation restrictions ($\leq$500 MHz). With that purpose, we studied the evolution of the ratio between the peak magnitude of the Fourier trans-



(a) 200-MHz FPT (b) 10-GHz FPT

Figure 3.16: Ratio between the peak magnitude of the field's Fourier transform in X and Y polarization versus the mixing angle. Two FPT center frequencies are evaluated, 200 MHz and 10 GHz, for 31 DGD cases ranging from 0 to 30 ps. The insets conceptually illustrate the alignment of the FPT with respect to the reference axis which, in this case, is arbitrarily chosen as the X polarization.

form in both polarizations (X and Y) against the mixing angle. This is performed for two different FPT center frequencies (200 MHz and 10 GHz) in the absence of data signal, and for 31 DGD cases (from 0 to 30 ps). The process is: initially, the FPT power is equally split between both polarization, which represents the worst-case mixing; then, certain degree of DGD is induced; and finally, the polarization rotation angle is swept and the ratio evaluated. Figure 3.16 presents the results. For 200 MHz and center frequencies in the same order of magnitude, PMD has hardly any effect in the ratio. This indicates that the angle at which the maximum peak magnitude is found, unequivocally tells the polarization mixing angle, thereby allowing for its pre-compensation before regular equalization. On the contrary, the interference of GHz-range FPTs appears highly sensitive to most of the considered DGD values; such that, from $\approx$15 ps on, the ratio ratio profile flattens out, preventing peak detection and thus reliable polarization de-rotation.

The use of one $\leq$500-MHz FPT if further justified as a means of realizing coarse estimation of the polarization mixing angle, accelerating the convergence speed in blind adaptive equalizers, and so reducing the required training time that partly determines their computational complexity. In addition, most of the FPT power will converge to one polarization state,

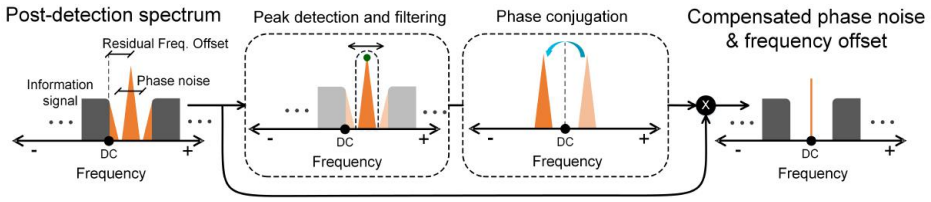


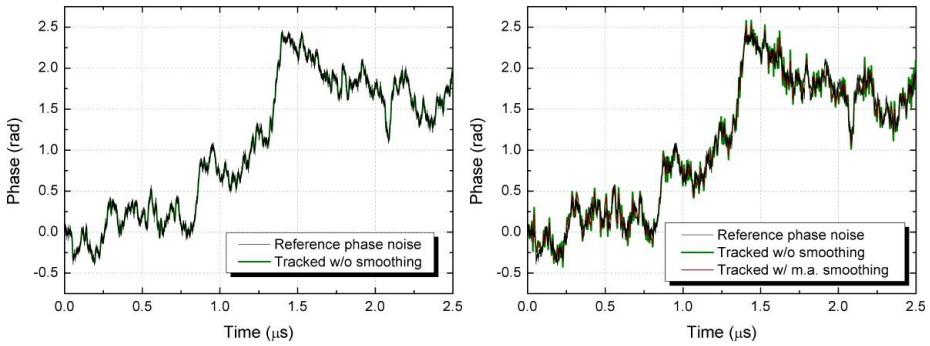
Figure 3.17: Conceptual schematic of 0-Hz FPT processing to compensate for residual frequency offset and phase noise.

which leads to higher pilot-to-noise ratios and therefore more precision in the next phase noise and frequency offset estimation process

Phase noise and frequency offset compensation : Figure 3.17 depicts the processing stages of the algorithm under discussion, where all computations preceding the correction of the estimated phase (i.e. peak detection, filtering, phase conjugation and tone reconstructions) are exclusively carried out on the polarization state containing the largest fraction of the FPT power. For simplicity, the example in Figure 3.17 concerns an FPT initially located at DC.

After Fourier transformation of the field, the maximum peak of the magnitude is detected. Maintaining frequency-domain operation, filtering is then applied centered around this spectral component to separate the FPT from the signal. The filter parameters and its type are selected based on empirical performance evaluation. In this investigation, rectangular or Gaussian profile is used, whose typical bandwidth ranges from 15 to 30 MHz primarily depending on the SNR level. At this point, the FPT is assumed perfectly isolated, and so the time-varying phase distortions are extracted by directly reading the signal argument. The measurement noise in the output phase vector may be attenuated through various smoothing tools, from which statistical techniques such as particle filtering [165], or extended Kalman filtering incorporating the laser's dynamics [166] have demonstrated remarkable performance. Finally, the equivalent phase-conjugated tone is reconstructed in the time domain, and multiplied sample-by-sample with the full detected C-MultiCAP signal. If the FPT is not initially inserted at DC, a simple correction factor corresponding to the default center frequency must be included in the estimated phase vector.

Figure 3.18 shows a simulated example of the FPT tracking performance



(a) 35-dB SNR (b) 15-dB SNR

Figure 3.18: Results of an FPT-based phase-noise tracking simulation at -15-dB PSR, and 35 and 15 dB of SNR. The reference vector is generated through a random walk process with Gaussian-distributed transitions exhibiting zero mean and 200-kHz equivalent variance.

at -15-dB PSR, and 35 and 15 dB of SNR (adjusted through AWGN loading). The data signal is a 4-level 6-band 2D-C-MultiCAP with 400-MHz guard-band for FPT allocation, β_b=0.1 and γ_b=0 $\forall b$. The isolating filter is Gaussian with 20-MHz bandwidth, and the smoothing technique is a simple moving average with ≈10-ns window. Frequency offset is not simulated in order to facilitate the visualization of the phase fluctuations, which are generated by means of a random walk process with Gaussian-distributed steps presenting zero mean and 200-kHz equivalent variance. The effectiveness of the FPT for phase-noise tracking is observed, as well as the apparent accuracy improvement that a simple moving-average smoothing may provide.

3.2.2.3 Blind channel equalization

Succeeding FPT processing and timing recovery, the output sequences from CAP matched filtering are decimated prior to blind channel equalization; generally entrusted to reverse all remaining intra- and inter-dimensional linear distortions. For instance in single-carrier transmissions, the joint compensation of time, polarization, and quadrature coupling (e.g. ISI, polarization mixing, and CD) is a well-known procedure, typically performed through a set of four complex-valued adaptive filters disposed in a 2×2×time MIMO *butterfly* structure [85], and possibly governed by different updating rules and error functions. In C-MultiCAP, however, the di-

mensionality is variable and MN_b times higher than single-carrier's, bringing the processing load well beyond pragmatic limits. Decreasing such complexity first requires identifying the least relevant coupling mechanisms, thus laying the ground for selectively grouping the independent sequences into lower-dimensional structures which can be later equalized in a simpler parallel fashion. In particular, we make the following observations:

- Because the residual frequency offset and phase noise are compensated beforehand, the IBI within each coherent-dimension will hardly degrade the performance. Inaccuracies in the estimation of the phase vector will lessen the validity of this consideration, nonetheless, we note that the magnitude of the deviation does not exceed 1% in normalized mean square error (NMSE) sense for >10-dB SNR under optimum PSR configurations. For instance, the smoothed curve in Figure 3.18b presents ≈0.5% deviation. In the presence of polarization and quadrature coupling, cross-dimensional IBI is conveniently reduced to its minimum by making all coherent-dimensions carry parameter-wise MultiCAP replicas. As advanced in Subsection 3.2.1.3, this arrangement assumes that the end-to-end filtering degradation is similar in the four coherent-dimensions, so that the same structure optimally applies; and the desire to exploit all operative spectral regions for SE maximization.

- Finer spectral partitioning makes the residual CD (after bulk compensation) and other filtering effects asymptotically equivalent to complex-valued scaling; such that, in the convergence limit, the ISI-driven coupling among CAP dimensions becomes suppressed. This results in CD and filtering impairments which only manifest as quadrature mixing from the bands' perspective.

- Imperfect FPT-based estimation of the polarization mixing angle, and the required PMD tolerance margins[18] compel to ≥2×2×time-dimensional MIMO equalization.

In short, CAP bands and dimensions are, to a large extent, exclusively coupled through quadrature and polarization mixing. This statement holds

[18] With average symbol rates per band as low as units of GHz, the mean DGD after 1000-km SSMF propagation with a characteristic PMD parameter of 0.1 ps/$\sqrt{\text{km}}$ represents less than 2% of the symbol period. However, tolerance margins recommended for high-speed MAN (see footnote 3) increase this fraction to approximately 20%-25% of the symbol period.

upon the assumption that the phase noise and frequency offset are neutralized, and that per-band filtering and CD degradation are well approximated by a flat complex-valued scaling. In such case, and given the imposed C-MultiCAP's structural restrictions, the joint equalizer can be simplified without risking performance by trimming off its dimensionality (i) M-fold, enabled through parallel processing per band; and (ii) N_b-fold, enabled through parallel processing per CAP dimension. Accordingly, the final equalization approach comprises a maximum of MN_b independent complex-valued 2×2×time-MIMO adaptive filter banks, for what the efficient and well-understood algorithms commonly employed in DP QAM single-carrier transmissions can be utilized. For facilitating the understanding, the reader is referred to Figure 3.19; a conceptual graphic showing the distribution of all available dimensions in C-MultiCAP systems, as well as their interaction before and after equalization under the above-mentioned assumptions and structural restrictions.

Comments on collective equalizer : besides decreasing the maximum number of jointly processed dimensions, this parallel approach also provides modular flexibility and diversity, then opening for strategic resource utilization through selective equalization; as well as the implementation of cooperative modes of operation with different purposes. In this investigation, three modes are devised:

- *Share :* channel estimation is performed on a single CAP dimension and the outcome is applied to itself and the remaining ones. Because the adaptive algorithm only needs to be run once per band, the complexity is minimized.
- *Average :* the channel is individually estimated for ≥2 CAP dimensions, the impulse responses are averaged for partial suppression of intra-band noise, and finally the resultant response is equally applied to all CAP dimensions. Intermediate compromise between complexity and performance.
- *Independent :* no cooperation between CAP dimensions. Linear impairments are individually corrected for each of them. Best performance and maximum complexity.

In addition, because the proposed method is not fundamentally limited by the number of bands or CAP dimensions, the system is potentially scalable to any M and N_b conditional upon satisfying the two core assumptions.

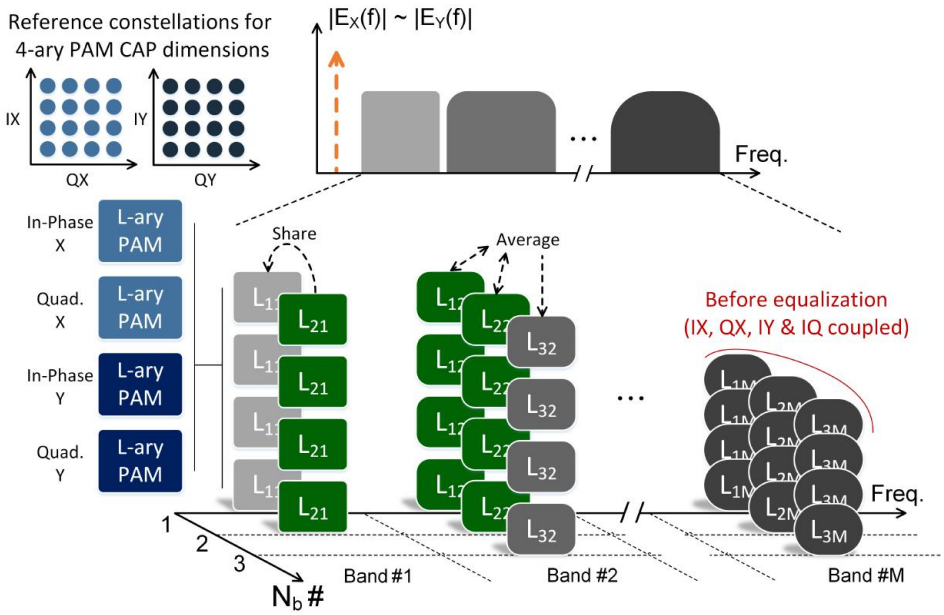


Figure 3.19: Conceptual distribution of the available dimensions in C-MultiCAP and their relation after equalization (Band #1 and Band #2), with all blocks clearly separated indicating orthogonality; and before equalization (Band #M), with the blocks vertically overlapped indicating coupling among coherent-dimensions only. On the top right, equivalent collapsed spectrum of the conceptual diagram; on the left-most side, the legend for each CAP dimension per band; and on top left, reference constellations for 4-ary PAM CAP dimensions (color-matched with the legend). The workflow of *Share* and *Average* equalization modes is also sketched, where CAP dimensions actively employed in the estimation of the equivalent channel response are colored in green, whereas dummy dimensions are colored in grey.

Comments on individual equalizers : blind compensation of the linear impairments per CAP dimension is therefore simplified to resolving polarization-, quadrature-, and time-wise coupling. To carry that out, we build two complex-valued sequences by assembling the two independent PAM signals of each polarization component, and treat them as the two complex-valued sequences of an equivalent DP QAM single-carrier transmission undergoing similar impairments. By doing so, we likewise make available these aggregate structures for their processing with mature tech-

niques such as CMA, RDE, or their decision-directed variants [84–86, 164]; whose proved simplicity and robustness when applied to blind equalization of QAM-like signals has fostered their pervasive acceptance in coherent optical communications. The interpretation of each CAP dimension as an individual DP QAM entity is grounded on the following observations:

- The amplitude modulation order is the same for all sequences within each CAP dimension; resulting in ensemble constellations outlining square-QAM-like distributions. Note that, however, the order is allowed to change among CAP dimensions (indicated by the different subscripts in Figure 3.19).
- Identical matched-filter pairs are used per CAP dimension across all coherent-dimensions. This configuration keeps 1:1 coupling proportion between polarization components and also between quadratures; and makes the above-mentioned square-QAM-like constellations remain valid as reference without requiring their modification to consider ISI effects (inherently induced when CAP filters are not matched).

In terms of implementation, certain simplification and performance increase can be obtained by contextualizing the above-mentioned algorithms to C-MultiCAP's DSP chain (see Figure 3.14b):

- After coarse compensation of the polarization mixing angle as well as the phase noise and frequency offset, decision-directed error functions can be directly applied on the equalizer's input. A reduction of the computational load and time is then possible by skipping the pre-convergence period, while improving the channel estimation accuracy by incorporating phase sensitivity.
- Due to precise matched filtering ensured by the prior FPT processing, symbol-spaced operation (in contrast to fractionally-spaced) is possible with negligible performance difference; further reducing the processing complexity if desired.

Simulation example : so as to validate the good functioning of the proposed equalization approach, we simulate the coherent reception of a 6-band 2D-C-MultiCAP system, operating at 20 dB SNR per band. The bands run at 4 GBd, and feature flat 4-ary bit-loading in all dimensions, flat power-loading, and β_b=0.1 plus γ_b=0 $\forall b$. A guard-band of 400 MHz is

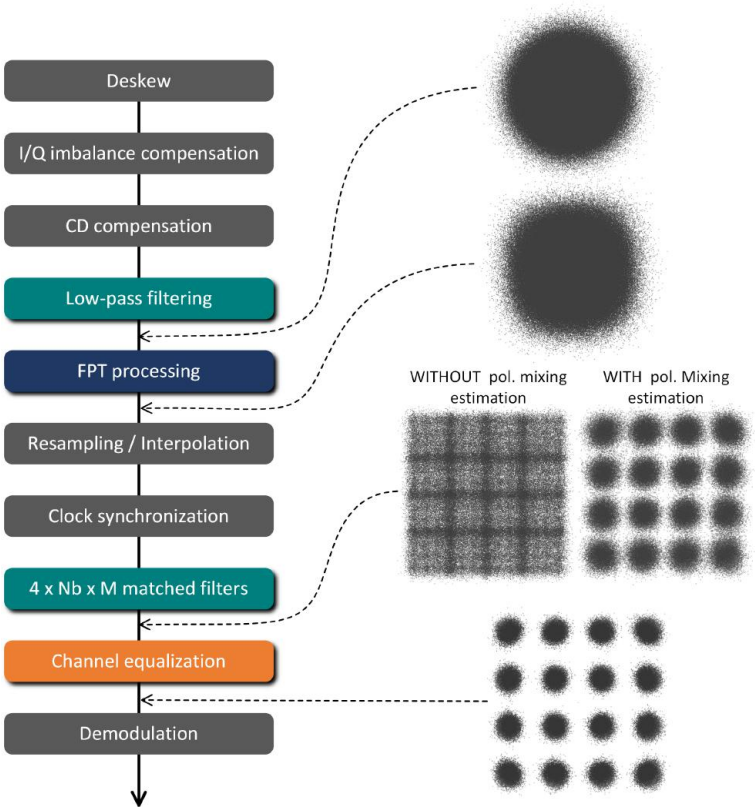


Figure 3.20: Evolution of the constellation diagram along the DSP chain for a simulation of 6-band 2D-C-MultiCAP detection. The constellations are all extracted from the X polarization of a single CAP dimension of the second band on the lowest-frequency side.

allocated for the FPT, whose PSR is adjusted for optimum performance. The FPT is employed for coarse polarization de-rotation, as well as for compensating 200-kHz variance phase noise (generated as indicated in Figure 3.18) and 200-MHz residual frequency offset. Concerning the individual equalizer, 4 filters with 10 symbol-spaced taps each are disposed in *butterfly* structure and iteratively updated by means of standard decision-directed algorithm [85] without pre-convergence interval. The equalizer is meant to compensate for 20-km residual CD (SSMF considered) and polarization mixing. Low-pass filtering effects were not simulated. The collective equalizer is set at *Independent* operation.

Demodulation below BER=10^{-3} (zero counted) was observed for all

bands over 30000 symbols per band. Figure 3.20 shows the constellations of the X polarization of one CAP dimension of the second band from the lowest-frequency side at different stages of the DSP chain. We selected this allocation for simplicity, as all constellations in all bands were alike in the absence of filtering effects. It is subtly appreciated how, after FPT processing, the raw constellation tends to square; hinting the satisfactory correction of the frequency offset. The constellations diagrams following matched filtering evince the usefulness of coarse polarization mixing estimation for facilitating the tap-updating convergence towards the reference 16-ary QAM constellation in the subsequent equalization stage; whose output confirms the feasibility of the proposed RX DSP scheme for its application in this concrete scenario.

3.3 Experimental demonstration

We have seen that C-MultiCAP systems, as conceived in this investigation, rely on a set of assumptions and generation restraints for enabling digital impairment compensation. Here summarize the most relevant ones:

- If frequency walk-off effect is disregarded, a single FPT can be used to compensate the whole signal for all time-varying phase fluctuations.
- If sub-bands are sufficiently narrow, the degradation caused by filtering effects and CD simplifies to complex-valued scaling, and thus fading out the coupling among CAP dimensions within the same band.
- If the post-detection residual frequency offset is ideally counteracted, so is the IBI. This leads to the complete isolation of all bands and CAP dimensions.
- If all coherent-dimensions carry parameter-wise MultiCAP replicas and use the same matched filters, the separate equalization of all CAP dimensions is possible by means of the standard algorithms employed in DP QAM single-carrier systems. This is also conditional upon satisfying the two previous points.

Imperfect compliance unequivocally ends in performance degradation or even total inability to recover the data depending on the deviation degree. Therefore, verifying the plausibility of C-MultiCAP for actual use in coherent communications first requires consolidating the soundness of these points in the intended experimental conditions. In our case, MANs are the

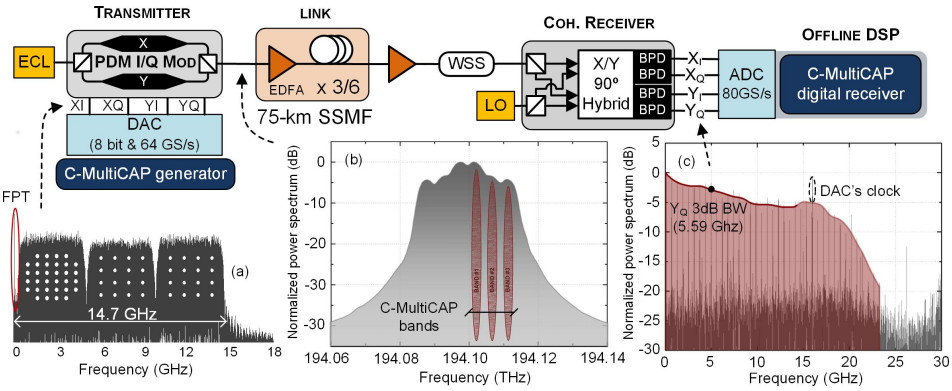


Figure 3.21: Schematic of the experimental setup for C-MultiCAP transmission. Insets: digital spectrum of the in-phase X polarization component of the 200-Gbps net rate case of study (a), its respective optical spectrum (b), and the post-digitization end-to-end system's frequency response from one of the input ports to the DSO.

targeted scenario, a network segment whose requirements and constraints naturally harmonize the potential features of C-MultiCAP. In this direction, we conducted the first proof-of-concept transmission of two high-rate C-MultiCAP configurations over a straight MAN-reach EDFA-only SSMF P2P link, employing just standard 25-GHz electro-optical components, and without nonlinearity compensation. The performance is evaluated in terms of BER for both B2B and transmission.

3.3.1 Testbed

Figure 3.21 shows the schematic of the experimental setup. The three insets on the bottom part show, on the left, the digital spectrum of the in-phase component on the X polarization for one C-MultiCAP case of study. Note the effect of power-loading targeting BER equalization among bands. The middle inset figure depicts the optical spectrum of the same case of study before the transmission link. And the right-most inset figure, which shows the end-to-end frequency response of the system at one of the RX ports after digitization.

At the transmitter, a 1544.5-nm ECL with ≈100-kHz linewidth is used as the light source. The output of the ECL is modulated by a DP I/Q modulator with a 3-dB bandwidth of 23 GHz. The electrical inputs to the modulator are pre-amplified with a set of four linear broadband amplifiers.

For data signal generation, one 4-port 64-GS/s DAC with 8-bit nominal resolution and 3-dB bandwidth of ≈12 GHz is used. The TX DSP consists of a C-MultiCAP signal generator as detailed in Subsection 3.2.1.3. The detailed description of two cases of study is included at the end. Linear M-shape digital pre-distortion is implemented to partly counteract the transmitter's low-pass attenuation, and ≈95% *clipping* factor is used to mitigate the quantization noise.

The output from the optical modulator is boosted with an EDFA and then launched into the fiber link for transmission. This link consists of 3 or 6 spans of ≈75 km of SSMF resulting in total transmission distances of 225 km and 450 km respectively. The propagation attenuation is compensated after every span employing EDFA-only amplification stages with ≈6 dB of noise figure. In the receiver side the signal is pre-amplified, filtered with a reconfigurable wavelength-selective switch, and subsequently detected in a standard polarization-diversity coherent receiver.

The integrated coherent receiver presented a 3-dB opto-electrical bandwidth of ≈22 GHz. One 100-kHz ECL is used as LO. After coherent mixing, the four outputs from the balanced photodiodes are digitized for offline processing by an 80-GS/s, 32-GHz bandwidth DSO. The RX DSP follows the steps described in Subsection 3.2.2. The demodulation is supported with k-means clustering as justified in Subsection 3.2.1, under the discussion on CAP's eye pattern.

C-MultiCAP generation : Two different sets of parameters are studied. Both are based on 2D-MultiCAP, where the original 2D-CAP technique (see 3.18) is employed for band generation and multiplexing[19]:

- Targeting 200 Gbps net rate, 3-band system with 32-, 16- and 16-ary QAM respectively[20] at 4.27 GBd each (221.88 Gbps gross rate), roll-off factor of 0.1 (15 taps per CAP filter) and 12% of the symbol

[19]In conventional 2D-CAP, both dimensions are typically paired to conform QAM constellation, which are not necessarily square (e.g. 32-ary QAM). This possible mutual dependence slightly complicates symbol generation and demodulation; nevertheless, it also serves for finer rate adaptation without opposing to the required assumptions or structural constraints to enable C-MultiCAP detection. Herein, adhering to this convention, we refer to the joint constellation traced by the 2 CAP dimensions within each band of each coherent-dimension as a K-ary QAM. This should not to be confused with the square-QAM-like constellation described by the same CAP dimension across the coherent-dimensions, which serves as reference for equalization (see Subsection 3.2.2.3)

[20]Bit-loading is initially adjusted for optimum BER performance in B2B configuration and left invariant throughout the experiment.

rate as inter-band spacing for all bands. This relaxed spectral distribution is chosen so as to minimize possible IBI while maintaining the entire signal within only 5-dB low-pass attenuation (see Figure 3.21). The electrical modulation bandwidth is ≈14.7 GHz, resulting in a total spectral occupancy after optical up-conversion of 29.4 GHz. This numbers yield ≈7.5 b/s/Hz of gross SE considering the actual occupied bandwidth, and ≈5.92 b/s/Hz gross SE resorting to the standard 37.5-GHz frequency slot.

- Targeting maximum throughput, 336 Gbps gross rate is attained with 5 bands of 4 GBd each using 32-, 32-, 16-, 16- and 8-ary QAM respectively. The roll-off factor is maintained in 0.1 (15 taps per CAP filter), however, in this case the inter-band spacing is necessarily reduced in order to squeeze the signal within the end-to-end 20-dB bandwidth. Accordingly, the resultant electrical modulation bandwidth is ≈21.5-GHz (≈15-dB low-pass attenuation at this point); translating into 7.81 b/s/Hz gross SE within the actual occupied optical spectrum (43 GHz), and 6.72 b/s/Hz gross SE in the standard 50-GHz frequency slot.

In both schemes, sub-band power-loading is implemented to minimize uncoded BER performance according to the end-to-end channel response. Concerning FEC, a hard-decision 7%-overhead product code with shortened BCH components is considered [167], achieving post-FEC BER of 10^{-15} at a pre-FEC of $\leq 4.4 \cdot 10^{-3}$ contemplating iterative decoding [168]. The reader is referred to Subsection 3.2.1.3 for detailed description of the FPT properties and insertion methodology.

3.3.2 Results

The discussion on the results is subdivided into (i) optical B2B, and (ii) Transmission.

3.3.2.1 Optical back-to-back

Figure 3.22 shows the BER performance as a function of OSNR for both C-MultiCAP generation cases in B2B configuration. Theoretical BER curves of equal-rate QAM signals are included for reference, where gray mapping and AWGN-only channel is considered for their calculation. Every mea-

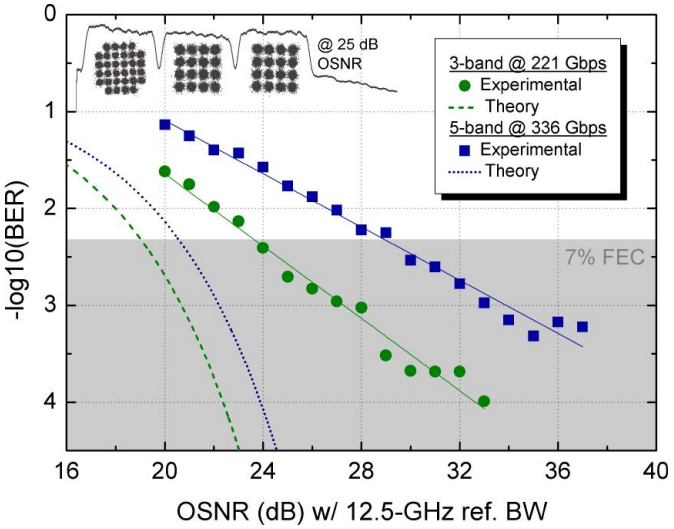


Figure 3.22: BER as a function of OSNR for optical back-to-back measurements. Inset: one instance of the recovered constellations in the in-phase component of the X polarization at 25-dB OSNR for the 221.88-Gbps gross rate case.

sured BER value is obtained from one sequence of ≈15000 symbols[21]. Performance below the FEC threshold is attained for both schemes, proving successful demodulation in the linear regime of operation.

The receiver sensitivity approximates 24-dB OSNR for 221-Gbps, and around 29 dB for 336 Gbps; which stands for 5 dB and 8 dB of implementation penalty respectively at the FEC BER threshold. The relative degradation difference of 3 dB is owed to two main factors. First, the bands in the 336-Gbps case are packed denser together, leading to IBI increase. And second, the outer part of the signal in the 336-Gbps case undergoes strong non-flat filtering; resulting in ISI-driven coupling between CAP dimensions in the highest-frequency bands. Both impairments partially break two of the fundamental assumptions that assure correct equalization, and therefore, strong degradation is understandable. Likewise, we estimate that other effects, such as PMEPR or partial noise correlation in the 336-Gbps case, are the reason for the slight discrepancy observed between both slopes; which is also believed to be the reason why both schemes exhibit disparate

[21] This is equivalent to $15000 \cdot 4 \cdot \frac{5+4+4}{3} = 259980$ bits in the 221-Gbps generation scheme and $15000 \cdot 4 \cdot \frac{5+5+4+4+3}{5} = 252000$ bits in the 336-Gbps one.

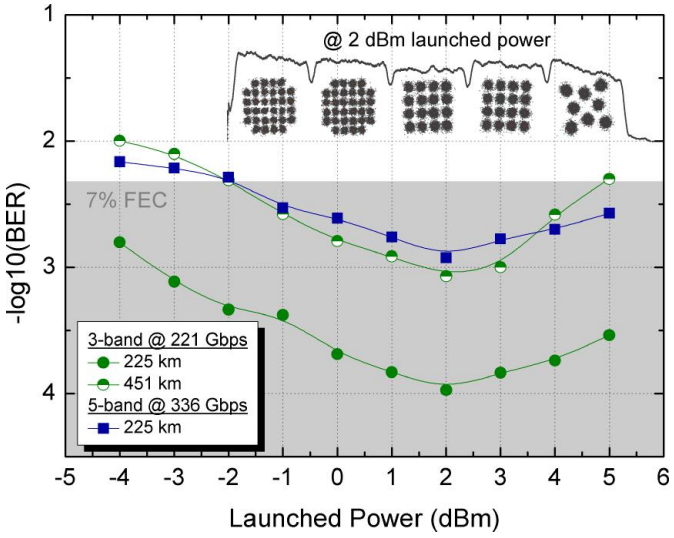


Figure 3.23: BER as a function of optical launched power for the two considered C-MultiCAP signals after 225-km and 450-km EDFA-only uncompensated SSMF transmission. Inset: one instance of the recovered constellations in the in-phase component of the X polarization at 2-dBm launched power for the 336-Gbps gross rate case.

error-floors ($\approx 6 \cdot 10^4$ for 336 Gbps, while not apparent error stabilization is observed for 221 Gbps down to $\approx 1 \cdot 10^4$).

Regarding the 5-dB base penalty of the 221-Gbps configuration, we simply attribute it to a general breach of compliance with the conditions for faultless equalization in combination with quantization noise. A clear violation example is visible in the inset of Figure 3.22, where even the lowest-frequency bands are non-flatly attenuated. These results and observations evince the need to (i) reduce the spectral occupancy per band below ≈4-GHz, to further mitigate non-flat filtering at the expense of processing complexity (see 3.23); (ii) trade SE for broader inter-band frequency guard interval, thus reducing IBI; and/or (iii) perfect the equalization technique, perhaps multiple stage equalization with alternative cross-dimensional grouping, which also requires extra processing power.

3.3.2.2 Transmission

Figure 3.23 presents the BER as a function of launched power. For 221 Gbps, the performance is evaluated after 225 km and 451 km of trans-

mission. For 336 Gbps, the BER results are only addressed for the 225-km transmission, since successful demodulation was not achieved in the longest link with the considered FEC. One example of recovered constellations at 2 dBm for the 336-Gbps case is shown for reference. BER values are each obtained from individual sequences of ≈15000 symbols (see footnote 21).

For 221 Gbps, transmission below the FEC BER threshold is observed for the full range of probed launched powers after 225-km transmission; and between -2 dBm and 5 dBm after 451 km. The non-linear threshold (best performance) is measured at 2 dBm for both propagation distances, a reasonable value for single-carrier systems that hints an acceptable level of PMEPR partly controlled with *clipping* technique. Interestingly, the transmission penalty between 225 km and 451 km at BER=10^{-3} equals ≈5 dB on the linear regime, approximately 2 dB higher than expected since we are just doubling the propagation distance[22]. This can be explained by the greater group delay that CD induces for longer uncompensated links, which increasingly worsens the accuracy of the FPT-based phase noise estimation for the highest frequency components (note that the FPT is allocated near DC). Imprecise phase noise estimation gives rise to an error accumulation along the subsequent DSP blocks (see Figure 3.14), not only translating in time-varying quadrature coupling but also imperfect equalization. This penalty manifests the importance of performing as exact carrier recovery as possible at the begining of the DSP in C-MultiCAP systems.

Concerning power margins, 4 dB and 3 dB are observed in the 451 km case for the linear and nonlinear regimes respectively; where the steeper non-linear slope as compared to 225-km transmission is simply owed to the higher accumulation of nonlinearities. In this context, despite the ample margins seen after 225 km of propagation plus the B2B results invite to think that the 336-Gbps system will also widely fit below the FEC threshold, up to 5 dB of penalty is measured at BER≈10^{-3}. This is consequence of the broader bandwidth aggravating the CD-induced walk-off, and again suggesting the consideration of an improved method for carrier estimation. One option consists of spreading 2 or more FPTs across the spectrum leveraging inter-band spacing (see Appendix C). The non-linear threshold for 336 Gbps is also found at 2 dBm input power, allowing for 4 dB power margin allocation (-2 dBm sensitivity).

[22]The OSNR over 12.5 GHz bandwidth around 1550 nm of wavelength is approximated by: $\mathrm{OSNR}|_{\mathrm{dB}} = 58 + P_{IN}|_{\mathrm{dBm}} - \mathrm{NF}|_{\mathrm{dB}} - \alpha L - 10\log N$, where P_{IN} is the launched power, NF is the equivalent noise figure of the lumped amplifier (EDFA in our case), α is the fiber's attenuation coefficient in dB/km, L is the span length, and N is the number of spans.

Overall, transmission results confirm B2B's in the need to perfect both C-MultiCAP parameters and RX DSP to guarantee robust transmission in coherent MAN scenarios. Particularly in this last analysis, the imprecision of the FPT-based phase noise compensation has been revealed; consequence of CD decorrelating the phase vectors among spectral components.

3.4 Summary

The burgeoning role of content-delivery networks combines with an unprecedented traffic volume and diversity to prompt the ongoing evolution of MANs towards ≥400-Gbps all-optical reconfigurable mesh architectures. These infrastructures enable high order of transparent and flexible connectivity which, nevertheless, come along with important technical challenges including pronounced low-pass end-to-end frequency responses caused by manyfold OADM cascadability plus electronics, larger PMD, and extra OSNR degradation given the absence of intermediate regeneration. In this context, high-baudrate single-carrier approaches have given way to super-channels and flexible-rate spectral partitioning techniques, proving extraordinarily convenient for counteracting these limitations [13]. Many-carrier super-channels result in lower symbol rates, thus relaxing the stress on the electronics and lowering the susceptibility to PMD and CD; as well as they enable precise spectral shaping, which improves the signal robustness and the per-channel SE. Nevertheless, the required degree of parallelism raises proportionally with the technical benefits, which given the stringent techno-economical requisites in MAN [91], poses a limit in the achievable granularity for hardware-based implementations.

In order to alleviate this trade-off, digital de/multiplexing has been contemplated as an powerful complement to parallel optics. Leveraging TX/RX DSP, these techniques provide software-level control of the signal's properties per carrier, translating not only into much finer spectral granularity but also into easy-reconfigurable spectral shape and throughput; all of which naturally suits the requirements of flexible transponders. This is the scope this investigation belongs to, where a novel DSP-based multiplexing and modulation technique is proposed and experimentally demonstrated.

This investigation

In this chapter, C-MultiCAP has been introduced; a P2P DSP-based multiplexing and modulation scheme which brings together the multi-

dimensionality of CAP and the spectral flexibility of MultiCAP for application in coherent communications. First, the essentials concerning signal generation and detection are explained in detail. Initially, the mathematical fundamentals and implementation concerns in multi-dimensional carrierless multiplexing are covered. We have seen that raising the dimensionality causes a quadratic increase of the total number of taps required for de/multiplexing such dimensions and, nevertheless, does not improve the SE and worsens the PMEPR. This leaves aside this feature for, perhaps, implementing multi-dimensional modulation formats which can potentially improve the receiver sensitivity by $\leq$1 dB within pragmatic complexity. We then extend the analysis to multiple bands (MultiCAP) where it is found that the total number of taps increases quadratically with both the number of bands and the number of dimensions per band. Accordingly, a set of practical generation parameters is proposed; whereby a spectral granularity down to $\approx$14.3% of the total bandwidth is possible with two dimensions per band and $\approx$1.5 dB PMEPR margin with respect to 256-carrier DMT without jeopardizing SE. These findings base the final elaboration on C-MultiCAP's generation and detection. It is argued why intradyne detection of C-MultiCAP signals requires early carrier recovery in the RX DSP as well as identical generation parameters in the four coherent dimensions to enable blind equalization taking advantage of conventional algorithms (e.g. CMA or RDE). Based on this, C-MultiCAP's RX DSP is conceived and explained along with its most distinctive blocks; that is, early FPT-based coarse polarization mixing compensation plus carrier recovery, and channel estimation per CAP dimension.

Following the characterization of C-MultiCAP generation and detection, we perform an experimental proof-of-concept with the purpose of verifying the validity of the assumptions supporting the proposed RX DSP in the target scenarios; and given imperfect compliance, evaluate the deviation tolerance. Two different 2D-C-MultiCAP configurations are implemented: 3-band case at 221.88 Gbps occupying 29.4-GHz optical bandwidth, and 5-band case at 336 Gbps occupying 43-GHz optical bandwidth. The BER performance is measured for B2B and transmission, where 225-km and 450-km EDFA-only SSMF links are considered. The main conclusion is that, under this combination of experimental conditions and generation parameters, the partial fulfillment of the basal assumptions translates into notorious implementation penalties ($\approx$5 dB and $\approx$8 dB for 221 Gbps and 336 Gbps respectively) and transmission penalties ($\approx$5 dB after distance doubling for 221 Gbps).

The results indicate the need to perfect both C-MultiCAP parameters and RX DSP to guarantee robust transmission. Narrower bands and multiple stage equalization with complementary parallelization approach stand for possible options in this direction, although the entailed increase in processing complexity partially questions their feasibility. Alternatively, certain increase in performance is achievable if looser inter-band spacing is applied. Despite this measure automatically leads to broader total bandwidths, we have demonstrated ≈8 GHz and ≈7 GHz optical spectral occupancy margin with respect to the tightest-fit standard grid for the 3-band and 5-band cases respectively; permitting this more relaxed allocation of bands without sacrificing per-channel SE. FPT processing has proved satisfactory for both polarization mixing angle estimation and frequency offset compensation. Nonetheless, with regard to phase noise estimation accuracy, strong penalty has been measured owed to CD-induced walk-off effect for long distances and broad signal bandwidths; which definitely motivates the use of multiple FPTs strategically inserted across the spectrum, and perhaps exploiting the above-mentioned looser inter-band spacing. On the other hand, the potential spectral and rate adaptability of C-MultiCAP have been demonstrated with up to 336-Gbps line rate at 7.8 b/s/Hz gross SE in a system exhibiting less than 6-GHz end-to-end 3-dB bandwidth. In addition, the convenience of C-MultiCAP's reconfigurability has also been shown, allowing for >100-Gbps line rate increase in the same link by simply adjusting the TX/RX DSP. Finally, recovery and equalization of multiple correlated CAP dimensions (e.g. those bands featuring 32-ary QAM) has been presented for the first time in coherent systems, proving the plausibility of their joint exploitation for implementing power-efficient multi-dimensional modulation formats.

Future work

The present investigation conveys the following items:

- Optimization of generation parameters and RX DSP : ensuring robust signal recovery is necessary first stage, and in this direction, several aspects still remain uncovered. For instance, per-band bandwidth and inter-band spacing have shown determinant in the equalization performance, and thus realizing scenario-dependent optimization can be particularly beneficial. Additionally, while the effectiveness of FPT-based techniques has been demonstrated, great transmission penalty reduction can be obtained by upgrading the proposed scheme

to two or more spread tones. This counteracts the CD-driven decorrelation of the phase noise, resulting in accurate matched filtering and thus error propagation avoidance. Finally, data-aided equalization should be contemplated. Because active polarization rotation tracking is performed by the FPT (anyways requiring data-aided disambiguation), the extra overhead owed to the training sequence could be substantially small.

- Beyond 2D-C-MultiCAP : after ensuring robust signal demodulation, further features can be contemplated. Despite processing complexity and PMEPR are proven limiting factors in C-MultiCAP that call for careful control, investigating the pragmatic benefits of those affordable >2D system configurations may be of interest (e.g. multidimensional modulation formats). Cross-band modulation/correlation also represents a plausible alternative.

- Non-linear tolerance : besides the well-known capability of FPT conjugation for partial nonlinearity compensation [46], recent studies suggest that sub-carrier/sub-band multiplexing employing very low symbol rates (2-10 GBd) partly reduces the non-linear noise interference in uncompensated lumped-amplified transmission links [169, 170]. Both techniques are present in C-MultiCAP, encouraging an interesting demonstration and characterization.

Chapter 4

Short-haul networks

Client-side networks comprehend the most varied technological ecosystem in the global communications hierarchy; showing reaches ranging from few meters to tens of kilometers, bit-rates from Megas to hundreds of Gigas, and propagation channels from wireless to SSMF. In this network segment, seamless communication is conditional on the punctual preparation of market-driven standards enabling equipment interoperability. The concrete implementation agreements vary according to the technical requirements of the service to support (e.g. reach, throughput or latency) and the environment of application (e.g. enterprise, residential, datacenters or service-providers' networks), all in all giving room to a broad spectrum of functional alternatives. For example, with the vast majority of wired local-area networks (LANs) operating on Ethernet[1], InfiniBand[2] and Fiber Channel[3] standards, more than 40 unique transmission speeds are considered in the aggregate roadmap including currently deployable and prospective configurations [171–173].

Within wired solutions, fiber-optic systems are unquestioned in scenarios beyond 100-m transmission and 10-Gbps rates[4] [171]; at the same time that their penetration in the $\leq$100-m market[5] is progressively favored by attributes like electromagnetic interference immunity and lightness [171, 174, 175]. Accentuated by the relentless increase of throughput

[1] http://www.ethernetalliance.org/

[2] http://www.infinibandta.org/

[3] http://fibrechannel.org/

[4] Router-to-transport and router-to-router applications where $\geq$10-Gbps 2-km, 10-km and 40-km duplex SSMF are industry modes; or warehouse-scale computing with thousands of 25-Gbps and 50-Gbps servers.

[5] In-house connectivity, airborne/automotive datacom or supercomputing.

demand, the technological panorama appears predominantly fiber-based; evincing the impact of optical LANs' design on the characteristics of the entire network. Fundamentally driven by cost and power consumption, optical LANs are typically built from duplex transceiver units employing intensity-modulation (IM) and direct-detection (DD) and application-specific combinations of fiber types and frequency bands[6,7]. In terms of modulation and capacity scaling, serialization (as opposed to parallelization) is preferred for reasons relating to simplicity of implementation, cost and power consumption [179, 180]. This premise governs the evolution of data communication standards, where resource-to-performance ratio optimization (e.g. electrical bandwidth, power, physical space and cost) is prioritized over the incorporation of new dimensions or its extension (e.g. wavelength, space, phase or polarization). For instance, enabled by the progress in ASIC design and DSP, digital equalization and FEC are now possible within acceptable latency margins. This has spur the recent inclusion of AMF into the scope of ≥100-Gbps standardization processes in order to boost the SE[8]. Nevertheless, as line-rate demand exceeds the capabilities of the affordable electrical bandwidth[9], parallelization and multiplexing are inescapably resorted to; on the whole yielding the hybrid technological variety that current transceiver modules exhibit.

Particularly in DD systems, WDM and fiber-pair parallelization have been favored as techniques that do not require retaining the field's phase information after photodetection, hence relieving the system from the extra complexity that it entails. Unfortunately, sustainable scalability is not exempt from saturation either, as tolerable power consumption, cost and per-blade transceiver density are surpassed for certain line rates and be-

[6] 850-nm single-λ/fiber or 850-to-940-nm WDM for ≤500-m transmission over laser-optimized multimode fiber and wideband multimode fiber respectively [176]. WDM O-band for ≥100-m SSMF transmission [177].

[7] Technological synergy among different modules is pursued during standardization for facilitating common testing and deployment, as well as the use of economies of scale for additional reduction in costs and time to market [178].

[8] In the second-to-third quarter of 2015, the IEEE 802.3bs 400 GbE task force moved to baseline adoption for 500-m reach employing 50-GBd 4-ary PAM (100 Gbps serial) and for 2-km and 10-km reach employing 25-GBd 4-ary PAM (50 Gbps serial) [177]. The 100GBASE-KP4 standard developed by the IEEE P802.3bj 100 Gbps task force will be used on the backplane ($4 \times$ 25-GBd 4-ary PAM lanes) [181]. Concerning FEC, hard-decision Reed Solomon (544,514,10) has been agreed on, with estimated total latency ranging from 80 ns to 120 ns depending on the concrete chip architecture and pre-decoding BER threshold $\approx 10^{-6}$ [182]

[9] This is especially pronounced in sectors like hyperscale datacenters and service providers' networks [171].

yond [17]. For instance, an excessive number of wavelengths requires strict wavelength stability control and makes packaging challenging; while too many fiber-pairs would likewise suffer from packaging and fiber-transceiver interfacing issues. Over and above, C-band transmission is avoided in client-side telecom scenarios[10] despite the ≈40% lower attenuation/kilometer with respect to O-band due to CD, hence tightening the power budget margins. All the latter naturally put the research spotlight on the previously precluded phase diversity, possibly allowing for quadrature multiplexing and effective CD compensation in the digital domain. On this subject, self-coherent approaches have been long studied as a means of enabling full-field recovery in standard single-detector systems by exploiting the photodiodes' quadratic response [184]. Despite proving successful for full field recovery, signal-signal beating terms cause notorious electrical spectral inefficiency, encouraging the development of solutions that trade, without exception, computational load [185, 186] and/or structural complexity [187, 188]. Interestingly, polarization diversity has been intensively investigated to circumvent the limitations of self-coherent approaches, resulting in numerous methods that partly counteract the mentioned inefficiencies [189–192], or even match the SE of single-polarization coherent transmissions [193, 194]. Both as complementary or alternative dimension, polarization-diversity has the potential to double the spectral efficiency, improve the power efficiency or provide the transmission systems with extra functionality without requiring explicit phase recovery (see Section 4.1).

Being optical fiber a transmission medium where polarization mixing changes over time, continuous or periodic tracking of the received field's state of polarization (SOP) is primordial in implementations of practical value. Often leveraging DSP capabilities, recent proposals conduct SOP rotation compensation in the digital domain, where computational load and latency become the bargaining chip. Emphasized by the stringent criteria on those very same resources, clear cause is given for SOP manipulation through simple real Mueller calculus on inherently phase-insensitive Stokes vectors. Unlike Jones calculus, Stokes parameters allow for (i) unequivocal SOP characterization under any degree of polarization (ii) in absence of phase information. Initially meant for polarization monitoring, Stokes receivers and Mueller calculus have lately burst into the research activities of DD data communications; showing especially suitable for DSP simpli-

[10]Based upon China Unicon data, 60% of core router to core router applications are covered with 10-km links, and 80%-90% with 40-km links. For its part, 95% of router to transport applications are covered with 10-km links [183].

fication in comparison to analogous Jones-based approaches. Following this work line, we present and experimentally demonstrate a DSP-mediate technique for polarization-multiplexed incoherent IM/DD systems, allowing blind tracking of the entire Stokes space. This enables multiplexing beyond two SOPs, while maintaining the simplicity of Mueller calculus and making possible zero-overhead continuous channel estimation without requiring matrix inversions.

This chapter is subdivided into four blocks. First, this investigation is positioned with respect to published literature concerning PolMux in DD optical transmission systems; distinguishing between phase-sensitive self-homodyne (coherent), and full-incoherent approaches. Second, the proposed fourth-order PolMux system is thoroughly explained, including its mathematical formulation, design requirements for the transmitter and receiver, and description of the employed Stokes-based DSP for channel estimation and SOP demultiplexing. Third, the experimental validation is presented, paying special attention to the receiver sensitivity and the robustness of the polarization-tracking algorithm. Finally, the work is summarized, comprehending a motivation recap, main results and conclusions, and plausible future work lines.

4.1 Work positioning

Stokes-based transmission systems admit as diverse classification as the purpose of polarization-diversity does, encompassing from SE increase to DD-channel linearization and receiver sensitivity improvement. However, for simplicity and without loss of generality, we employ the field's phase preservation after photodetection as the organizational factor. Accordingly, in this section we present the most recent investigations on DD transmission systems operating on Stokes space for SOP tracking and rotation compensation categorized as incoherent or coherent. Our work is contrasted with the literature and the major contributions highlighted.

Incoherent

The absence of phase information and the wish for low processing complexity motivates the use of incoherent detection in scenarios where CD and PMD hardly influence the system's performance. Within this family of techniques, we subdivide the published work according to the domain/s-

pace where symbol design and detection is realized. An alternative perspective of this subdivision is that we differentiate between DD systems that operate on the Stokes space strictly for SOP rotation compensation; and those which, additionally, decide on the transmitted symbols directly on Stokes space. These options are referred to as *Jones-Stokes* and *Stokes* respectively.

Jones-Stokes

Simulation results:

- [195]: a blind DSP algorithm for SOP tracking and demultiplexing for on-off keying (OOK) dual-polarization signals is introduced and characterized. A standard Stokes analyzer [196] is employed in the receiver side. The simulations account for phase noise, amplified spontaneous emission (ASE) noise and random fiber birefringence (polarization-dependent loss (PDL) is disregarded). The BER results against different carrier-to-noise ratios per polarization show negligible sensitivity from digital demultiplexing even under very fast SOP fluctuations.

- [197]: the same author adapts the algorithm in [195] to support dual-polarization PAM signals by adjusting the decision thresholds of the SOP tracker. The system structure is preserved.

Experimental results:

- [198]: single-laser dual-polarization 4-ary PAM SSMF transmission is demonstrated using 1310-nm silicon photonic intensity modulator. Successful demodulation is reported at a maximum line-rate of 224 Gbps after 10-km link. The Stokes receiver is made up of a 2×4-90^{o} optical hybrid, a variable optical delay line and 6 photodiodes. The DSP algorithm achieves polarization demultiplexing plus ISI cancellation through a set of 4 multi-tap adaptive filters applied on the respective Stokes parameters. The updating rule is MMSE on 2 different error signals.

- [199]: Extending the their work in [198], single-laser dual-polarization 4-ary PAM plus 4-ary inter-polarization phase modulation (6 bit/symbol) is demonstrated in B2B configuration at a maximum gross rate of 350 Gbps (soft-decision 25% overhead FEC is con-

sidered). DSP comprises two stages, first multi-tap 4×4 MIMO processing inverts the induced polarization rotation and partially compensate the end-to-end ISI. The taps are updated using MMSE on training symbols prior to decision-directed mode. Second, 4 extra multi-tap adaptive filters (interpreted as one per Stokes parameter) remove the residual ISI.

Stokes

Symbol distribution design for optimum detection on the polarization space increases the modulation flexibility and, potentially, the receiver sensitivity at the expense of transmitter complexity and cumbersome equalization. This modulation scheme is known as polarization-shift keying (PolSK) or polarization modulation, and the initial investigations involving Stokes receivers date back to late 80's [200]. From that moment, a series of reference investigations by S. Betti et al. [201,202], S. Benedetto et al. [203–207], and M. Nazarathy et al. [208, 209] thoroughly explored this modulation format and its potential benefits from different perspectives; e.g. theoretical fundamentals, multilevel and multipower PolSK performance for different noise sources, optimum receiver structures, optimum constellations in SER sense, or the development of optimization frameworks. These results pave the way for today's understanding[11] which, seizing on the evolution of DSP, has given rise to newly-published 100-Gbps multilevel PolSK proposals exhibiting tolerable complexity [210]. Continuing the work in [211], in [210] the authors elaborate on the receiver structure and the blind DSP algorithm for SOP tracking and demultiplexing. Via simulations, the BER of up to 32-ary PolSK targeting 100-Gbps line-rate is assessed accounting for phase noise, ASE noise and random fiber birefringence (PDL is disregarded). Particularly for 16-ary PolSK, as low as ≈3-dB sensitivity penalty is attained with respect to single-polarization IM binary scheme at BER=10^{-5} despite the 400% greater SE.

Coherent

Stokes receivers have also been employed in conjunction with self-coherent detection for simplifying the linearization of the DD channel, thence making possible to retrieve the field's phase information. In addition to quadrature

[11] Note that the literature on PolSK extends well beyond the alluded work, however, we only considered fundamental investigations and/or those related to Stokes-based detection.

multiplexing, phase preservation permits digital manipulation for effective CD compensation, allowing for C-band transmission with ≈0.14 dB/km power-budget increase with respect to O-band in SSMF. The main idea is that, in the absence of polarization mixing or complete compensation, a simple linear combination of the Stokes parameters S_2 and S_3 bears the electrical field along one Jones vector's elements (e.g. X or Y) when an unmodulated carrier copy is transmitted on the orthogonal complementary. Due to the spectral superposition of data signal and optical carrier, the electrical SE is not diminished and the optical bandwidth is minimized (same SE as single-polarization coherent transmission). These are the main advantages with respect to polarization-diversity self-homodyne alternatives without Stokes detection [212]. Recent results include:

- Repeatered uncompensated 160-km SSMF transmission of OFDM with flat 16-ary and 64-ary QAM bit-loading is reported at 160-Gbps and 80-Gbps gross rate respectively [213]. Four months later, the same group demonstrated single-carrier 32-ary QAM at 62.5 Gbps transmitted over the same link configuration [214]. One month later from the latter, the case of OFDM with flat 16-ary QAM bit-loading in [213] was subsequently extended to WDM configuration for ≈1 Tbps gross-rate over repeatered uncompensated 480-km SSMF transmission [215]. All these experimental validations operated on the C-band and employed Stokes receiver.

- [212, 216, 217]: based on the above-mentioned investigations, the same group revisited the topic in a series of publications. These contain details on the Stokes-based DD channel linearization concept, analysis on PMD-induced performance degradation, DSP algorithm elaboration, as well as discussions on experimental results.

- [218]: DSP algorithms for PMD tolerance improvement when OFDM signal is employed are experimentally evaluated based on the simulation work in [219]. Under 1 dB Q-factor penalty is observed up to 10-ps DGD[12] compared to the original DSP chain in [213, 215].

[12]When the spectral occupancy of the polarization-multiplexed signal $\ll$ the central frequency, the impact of PMD is customary estimated from the average DGD; directly related to the propagation distance as $\langle \text{DGD} \rangle \approx 0.92 D_p \sqrt{L}$, where L is the reach in kilometers and D_p is a characteristic fiber parameter in the order of 0.1 ps/$\sqrt{\text{km}}$ for modern SSMF.

The technique presented in this chapter belongs to the incoherent *Jones-Stokes* group, and it goes beyond the state of the art by enabling up to quaternary (four-fold) polarization-multiplexed fiber transmission with low DSP complexity. The software does continuous and blind tracking of the full channel's Mueller matrix (in contrast to tracking only a particular vector projection onto one of the components of the Stokes-space basis [195]), making possible the compensation of very fast polarization mixing effects on any number of simultaneously transmitted SOPs constituting a non-singular ensemble. We exploit this capability to simultaneously transmit four independent data streams carried by four wavelengths with different SOPs each, and employ polarization as the selective factor for demultiplexing the carriers without employing wavelength-selective devices or phase information. Note that dual-polarization configuration for SE doubling shall be understood as a lower-dimensional particularization of the system under discussion, implicitly demonstrating its feasibility for such purpose as well.

In the next sections we elaborate on the proposed four-fold polarization-multiplexed subsystem including conceptual description, mathematical model, detailed explanation of the software employed for SOP tracking and demultiplexing, and experimental validation after 2-km SSMF at 100 Gbps (128-Gbps gross rate) and non-return-to-zero (NRZ) modulation.

Contributions of this work

- Original idea: developing a general PolMux technique for simultaneous transmission and demultiplexing of up to 4 disparate SOPs in DD communication systems. Proposal of such system for its use alongside WDM technology to perform wavelength demultiplexing leveraging polarization diversity, thereby avoiding the need for wavelength-selective devices.

- Experimental results: experimental validation for short-reach ($\leq$2 km) C-band DD transmission systems.

- Related first-author publications: [220] (post-deadline conference), [221] (invited journal) and [222] (invited conference).

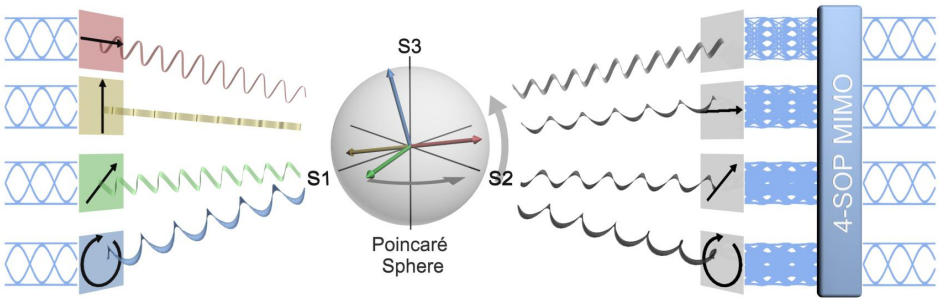


Figure 4.1: Conceptual scheme of the proposed quaternarty polarization-multiplexed data transmission DD system employing Stokes analyzer.

4.2 Quaternary polarization-multiplexed direct-detection system

The traditional approach to PolMux is to encode independent data streams into two linear and orthogonal polarization components (X and Y). The orthogonality avoids the interference between the polarized waves in the absence of polarization mixing or after compensation, thus leading to SE doubling. Nonetheless, it does not limit the multiplexing order to two. Herard and Lacourt [223] were the first to take advantage of this fact in an optical communication system, experimentally demonstrating a free-space IM/DD link with three linear SOPs and analog demultiplexing. This was followed by [224], with an extended theoretical analysis accounting for the impact of various factors on the system's performance including PMD and mode coupling. At that time, no more than three SOPs were simultaneously transmitted, and no algorithm for real-time channel estimation was presented. In this work, we demonstrate for the first time a quaternary polarization-multiplexed (4-SOP) transmission over a medium in which polarization mixing/rotation occurs, i.e. optical fiber. The consideration of circular SOPs enables forth-order multiplexing, and Stokes-based DSP permits blind and simple polarization rotation compensation without requiring the field's phase information.

4.2.1 Conceptual description

The proposed 4-SOP system is illustrated conceptually in Figure 4.1. Four independently IM optical signals modulate four different C-band carriers (distinguished by color) in four different SOPs (arrows on the left/trans-

mitter side) for later simultaneous transmission over SSMF, where polarization rotates over time [225] (indicated by the vectors in the Poincaré sphere). At the receiver, the field's intensities along SOPs *X*, *45°* and *right-circular (RC)* (or orthogonal counterparts) are independently detected, together with the instantaneous total intensity in a standard Stokes analyzer [196]. Subsequently, polarization rotation compensation, demultiplexing and demodulation (encompassed within *4-SOP MIMO*) are carried out in the digital domain.

4.2.2 Mathematical description

In matrix form, the noisy[13] system up to photodetection (inclusive) can be expressed as:

$$\mathbf{Out} = \mathbf{H_{sys}}\,\mathbf{In} + \mathbf{Noise}, \tag{4.1}$$

where $\mathbf{H_{sys}}$ is a $N \times N$ matrix, with N equal to the number of transmitted SOPs, which includes both the channel linear distortions and the predefined input-output SOP configuration (e.g. standard dual polarization). **In** and **Out** constitute the transmitted and received signals respectively. **In** is constructed from the N independently modulated intensity (real) waveforms. **Out** represents the observed system outputs after photodetection (real), and it results from the transformation of the transmitted intensity signals (**In**) through $\mathbf{H_{sys}}$. And **Noise** is the system's noise in matrix form.

At the receiver the distortions are compensated and the polarizations demultiplexed (using 4.1):

$$\mathbf{In^{rec}} = \mathbf{H_{sys}^{-1}}\,\mathbf{Out} = \mathbf{In} + \mathbf{H_{sys}^{-1}}\,\mathbf{Noise} = \mathbf{In} + \mathbf{W}, \tag{4.2}$$

Because, in our case, the necessary polarization diversity for SOP demultiplexing is obtained by detecting the field's intensity along various polarization orientations[14], the real transfer function of the proposed IM/DD

[13] Given that optical amplification is not considered in this application, the dominating noise is predominantly AWGN from the electrical domain (mostly from transimpedance amplifiers). In cases were optical amplification is employed (e.g. in some self-homodyne system proposals, see Section 4.1 under coherent group), the noise statistics becomes a mixture of the electrical AWGN and non-negligible squared ASE. It is then foreseeable, though not proved in this report, that the quadratic noise component will limit the maximum reach and system performance through accuracy degradation of the SOP tracking process, which always operates on intensities (Stokes parameters). This suggests using multilayer non-linear equalizers (MLP equalizer [226]) for prior SNR enhancement.

[14] Alternatively, some receivers attain polarization diversity by optically manipulating just orthogonal polarization components of the field (e.g. X and Y) prior to

system ($\mathbf{H_{sys}}$) can be readily calculated in two steps. First, the field transfer function is constructed:

$$\mathbf{G_N} = \begin{pmatrix} \vec{v}_{R,1} \\ \vdots \\ \vec{v}_{R,M} \end{pmatrix} \mathbf{H_{dist}} \begin{pmatrix} \vec{v}_{T,1}^{\top} & \cdots & \vec{v}_{T,N}^{\top} \end{pmatrix}, \tag{4.3}$$

where $\vec{v}_{T,j}$ and $\vec{v}_{R,i}$ are the 1×2 Jones vectors[15] of the jth and ith transmitted and detected SOPs respectively, and $\mathbf{H_{dist}}$ is the 2×2 Jones matrix that represents the polarization rotation undergone by the field along propagation over fiber. Note that number of individual receiver outputs (i.e. number of individual output photocurrents), M, is not necessarily equal to the number of multiplexed SOPs, N.

Second, the intensity (real) system matrix is obtained from $\mathbf{G_N}$ applying the Hadamard product operator:

$$\mathbf{H_{sys}} = \mathbf{G_N} \circ \mathbf{G_N^*} \tag{4.4}$$

In the absence of polarization mixing or after compensation ($\mathbf{H_{dist}} = I_2$), $\mathbf{H_{sys}}$ strictly describes the transfer function of the ideal polarization-multiplexed transmission system; herein referred to as *mapping matrix*. In this case, if $\vec{v}_{T,j} = \vec{v}_{R,i}\ \forall j = i$ and when linear SOPs are used, $\mathbf{H_{sys}}$ can be conveniently constructed by direct evaluation of Malus's law[16] [223].

It is apparent from 4.2 that $\mathbf{H_{sys}}$ must be invertible in order to make demultiplexing possible . Since the channel properties cannot be changed as part of the design procedure, the careful choice of $\vec{v}_{T,j}$ and $\vec{v}_{R,i}$ is the only method to assure $\det(\mathbf{H_{sys}}) \neq 0$. Therefore, by inserting Equation 4.3 into Equation 4.4 for different SOP configurations, the existence of $\mathbf{H_{sys}^{-1}}$ shall be evaluated under the assumption of perfect compensation of the channel distortions ($\mathbf{H_{dist}} = I_2$). This matrix (inverse of *mapping matrix*) is herein referred to as *demapping matrix*.

photodetection. Examples recently suggested for data communications can be found in [198, 214, 217, 227].

[15] Its mathematical expression is given by $\vec{v} = \begin{bmatrix} E_{0x} \cdot e^{i\phi_x} & E_{0y} \cdot e^{i\phi_y} \end{bmatrix}$, where E_0 and ϕ represent the complex amplitude and the absolute phase of the electromagnetic field along the polarization components X and Y. Patently, any Jones vector is fully described by no more than four independent variables (one complex number per dimension).

[16] The total intensity of a light beam going through a perfect polarizer is dependent on the incident angle θ as follows: $I_{out} = I_{in} \cos^2 \theta$

Stokes space : The four real Stokes parameters (S_{0-3}) are the elements of the Stokes vector ($\vec{S}$), satisfying $S_0^2 \geq S_1^2 + S_2^2 + S_3^2$ (see[17]) and unequivocally describing the polarization characteristics of the electromagnetic field at the desired time instant. Therefore, the explicit detection of any combination of field's information variables enabling the reconstruction of the Stokes parameters (e.g. coherent receiver, Stokes receiver or Stokes analyzer) suffices for assuring successful demultiplexing of any non-singular set of SOPs. Particularly in the considered DD system (described by 4.3 and 4.4), this avoids the need to determine $\vec{v}_{R,i}$, now fixed according to the concrete receiver structure and constituting a full polarization-diversity basis (see Section 4.2.4). By fixing $\vec{v}_{R,i}$ to such a basis, corroborating the existence of *demapping matrix* can be simplified to find those $\vec{v}_{T,j}$ that satisfy $\det(\mathbf{H_{sys}}) \neq 0$ for $\mathbf{H_{dist}} = I_2$ when $\vec{v}_{R,i}$ in 4.3 are replaced by the $\vec{v}_{T,j}$ themselves. This reduces the number of open variables to N (instead of $N+M$).

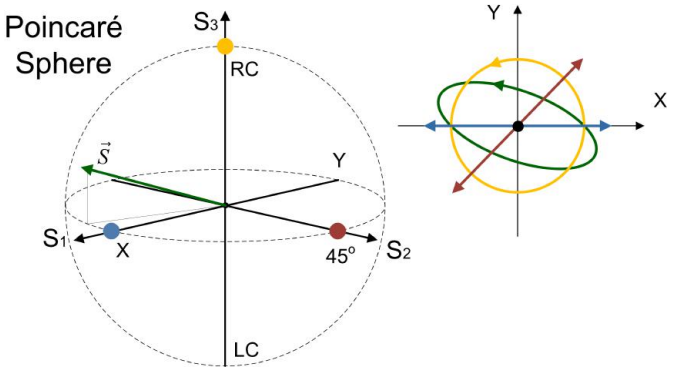


Figure 4.2: Poincaré sphere with highlighted SOPs (in color). Their projection on the $X \perp Y$ plane is included for reference (color matched).

For its part, the Stokes space is the four-dimensional vector space spanned by $\vec{S}$, often represented in a S_0-normalized three-dimensional Euclidean space with S_{1-3} as coordinate system for intuitive visualization and analysis (Poincaré formalism). Figure 4.2 shows the Poincaré sphere

[17]The inequality becomes equality when light is perfectly polarized, i.e. the degree of polarization ($\wp$) equals 1 [196]. Otherwise, we talk about partially polarized ($0 \leq \wp \leq 1$) or unpolarized ($\wp = 0$) light, which can be interpreted as unstable (noisy) and completely random SOPs over time respectively. When light is fully polarized, both Jones calculus (Jones vector matrix manipulation) and Mueller calculus (Stokes vector matrix manipulation) can be employed, otherwise Mueller calculus is mandatory.

with the coordinates of arbitrary SOPs, as well as their equivalent representation on the $X \perp Y$ plane (along the direction of propagation). It becomes visible that any SOP distortion maintaining the field's intensity unmodified translates to a simple real three-dimensional (3D)-rotation of the Stokes vector. This fact enables tracking of the SOP changes occurring during propagation over optical fiber employing cost-effective DD systems without compromising processing complexity.

4.2.3 Transmitter

Recalling the findings in [223], when linear SOPs are used as selective agents for signal de/multiplexing, the maximum achievable diversity equals three. This is intuitively explained from the reformulation of the Jones vector in footnote 15 as:

$$\vec{v} = \begin{bmatrix} |E_{0x}| \cdot \cos\alpha & |E_{0y}| \cdot \sin\alpha \cdot e^{i\delta} \end{bmatrix} \tag{4.5}$$

where $|E_{0x}|$ and $|E_{0y}|$ represent the magnitude[18] of the field along SOPs X and Y; α is the azimuthal rotation of the field's SOP (often referenced to the SOP X); and δ is the phase difference between SOPs X and Y (i.e. $\delta = \phi_y - \phi_x$) which equals 0 for linear SOPs. Two decisive conclusions are extracted from 4.5: (i) at each time instant, the superposition of any number of linear SOPs traces an equivalent composite linear SOP:

$$\vec{v}_{comp} = \sum \begin{bmatrix} |E_{0x}| \cdot \cos\alpha & |E_{0y}| \cdot \sin\alpha \end{bmatrix} = \begin{bmatrix} \sum |E_{0x}| \cdot \cos\alpha & \sum |E_{0y}| \cdot \sin\alpha \end{bmatrix} \tag{4.6}$$

and (ii) any linear SOP is unequivocally described by no more than three independent parameters (i.e. $|E_{0x}|$, $|E_{0y}|$ and α). An illustrative example is shown in Figure 4.3. Therefore, when more than three linear SOPs are superimposed, their mutually-independent intensities (e.g. data) cannot be recovered from the measured linear SOP (N independent inputs by 3 independent outputs). This leads to an impractical system with irreversible generation process and then, to the misconception that no more than three SOPs can be simultaneously transmitted and demultiplexed [223, 224].

In order to scale the reversible multiplexing order beyond three, the number of total independent parameters required to describe the composite SOP at all time instants must be strictly greater than three. Mathe-

[18] It is important to realize that the signed amplitude of the field is not available in incoherent DD systems, but rather the magnitude is detected.

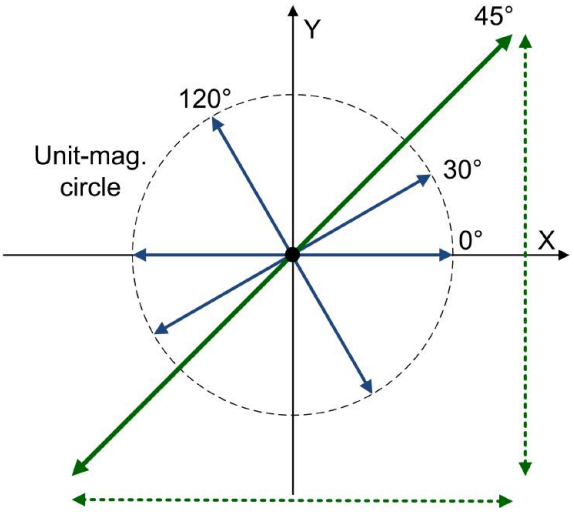


Figure 4.3: Example of superposition of three linear SOPs (*0°*, *30°* and *120°*, in blue solid) and the ensuing aggregate SOP (*45°*, in green solid). The direction of the electric field oscillations are displayed as double-ended arrows scaled by the magnitude of the corresponding SOP along the traced angle. The unit-magnitude circle (dashed black) is included for reference.

matically, this enables the design of input-output characteristic matrices (whose coefficients are associated with physically meaningful independent variables) defining consistent systems of more than three equations with unique solution (introduced in Section 4.2.2 as *mapping matrices* with non-zero determinant). Based on 4.6 it is apparent that, besides $|E_{0x}|$, $|E_{0y}|$ and α, the relative phase (δ) constitutes a fourth independent variable to fully describe the SOP of any polarized electromagnetic wave. In this work, we exploit this extra degree of freedom by encoding part of the information on a circular SOP, thus resulting in a composite SOP unequivocally described by (and no more than) four independent parameters. Particularly, the four transmitted Jones vectors ($\vec{v}_{T,j}$) employed in the experimental demonstration are X, Y, *45°* and LC.

$$\mathbf{Tx} = \begin{pmatrix} \vec{v}_{T,N} \\ \vdots \\ \vec{v}_{T,N} \end{pmatrix} = \begin{matrix} \\ X \\ 45^{\circ} \\ Y \\ LC \end{matrix} \begin{matrix} \begin{matrix} X & Y \end{matrix} \\ \begin{pmatrix} 1 & 0 \\ \cos\frac{\pi}{4} & \sin\frac{\pi}{4} \\ 0 & 1 \\ \cos\frac{\pi}{4} & i\cdot\sin\frac{\pi}{4} \end{pmatrix} \end{matrix} \tag{4.7}$$

As indicated in Section 4.2.2, when the use of a receiver structure offering full polarization-diversity is assumed, the existence of *demapping matrix* for a set of given $\vec{v}_{T,j}$ vectors can be evaluated from $\det(\mathbf{H_{sys}}) \neq 0$ by making $\vec{v}_{R,i}$ equal to 4.7 in the calculations (the actual $\vec{v}_{R,i}$ may be different, and thus the *mapping matrix* too):

$$\mathbf{H_{sys}} = \begin{pmatrix} 1.0 & 0.5 & 0.0 & 0.5 \\ 0.5 & 1.0 & 0.5 & 0.5 \\ 0.0 & 0.5 & 1.0 & 0.5 \\ 0.5 & 0.5 & 0.5 & 0.0 \end{pmatrix} \tag{4.8}$$

with $\det(\mathbf{H_{sys}}) = -0.25$ and inverse:

$$\mathbf{H_{sys}^{-1}} = \begin{pmatrix} 1.0 & -1.0 & 0.0 & 1.0 \\ -1.0 & 2.0 & -1.0 & 0.0 \\ 0.0 & -1.0 & 1.0 & 1.0 \\ 1.0 & 0.0 & 1.0 & -2.0 \end{pmatrix} \tag{4.9}$$

The above matrices prove that quaternary PolMux is possible in IM/DD systems by confirming the reversibility of one particular example of multiplexing scheme (4.7).

Practical considerations : It is conspicuous from 4.8 and 4.9 that the fact that multiplexing is reversible does not imply that the transmitted SOPs are uncoupled. Consequently, multilevel signals are formed at the output of all the photodiodes owed to the incoherent superposition of the $\vec{v}_{T,j}$-weighted intensities even without polarization rotation. This contrasts with orthogonal polarization multiplexing. Therefore, assuming that the receiver's electrical AWGN is the dominating noise source (see footnote 13) and the perfect channel estimation, the receiver sensitivity with respect to single-polarization 4-λ WDM with ideal optical filtering is mainly penalized by the SOP-demultiplexing process ($\mathbf{W}$ in 4.2 for $\mathbf{H_{dist}} = I_2$); which unavoidably couples the diverse independent noise contributions of each SOP according to $\mathbf{H_{sys}^{-1}}$. For instance, in N^{th}-order orthogonal multiplexing $\mathbf{H_{sys}^{-1}} = I_N$ after compensating channel distortions, evincing the mentioned signal degradation when compared with 4.9. Likewise, it becomes immediate from 4.9 that not all the channels undergo the same penalty, and that the fewer the number of coupled $\vec{v}_{T,j}$ are transmitted (either by imposing mutual orthogonality or by reducing the multiplexing order), the lower the sensitivity degradation will be. This in turn justifies the use of X and Y on the transmitted set of SOPs (see 4.7).

Despite this inherent degradation, $\det(\mathbf{H_{sys}}) \neq 0$ indicates that the set of SOPs defining $\mathbf{H_{sys}}$ can be successfully employed as demultiplexing selective agents of any other multiplexing technology, hence enabling the interchange of the respective de/multiplexing apparatus. This idea inspired the application of PolMux along side WDM in order to replace the wavelength-selective components, thus considerably relaxing the requirements on lasers' frequency stability, and on the power consumption due to the thermal control of the gratings.

4.2.4 Receiver

Random fluctuations of fiber's birefringence along the link cause a time-varying oscillation of the aggregate SOP (aSOP)[19]. In other words, fiber propagation[20] may transform (usually expressed as *rotation*) the transmitted SOP into any other state, which translates into polarization mixing after detection with a static polarization-diversity receiver[21]. Consequently, compensating such effect needs acting on all four independent parameters of the received SOP, and provided that phase information is lost after photodectection, a quaternary polarization-diversity receiver is necessary irrespective of the multiplexing order. This represents an important rule for the receiver design of polarization-multiplexed DD systems, being the 4-SOP setting the only 100% hardware efficient (4 independent inputs by 4 photodetectors[22]).

Hardware-wise, the simplest receiver structure comprises a one-to-four power splitter prior to a fixed set of polarizers, which provide the necessary polarization diversity between the various signal copies. Each of the branches ends with respective photodetectors for subsequent digitization and processing in the digital domain (see Section 4.2.5). The polarization filtering stage constitutes the only optical processing in the receiver, becoming the sole degree of freedom for field manipulation before acquisition. This makes the concrete $\vec{v}_{R,i}$ configuration (in relation to $\vec{v}_{T,j}$) have

[19] Composite SOP conformed by the superposition of SOPs described by $\vec{v}_{T,j}$.

[20] With the exception of polarization-maintaining fibers

[21] PDL is assumed negligible and $\langle \mathrm{DGD} \rangle \lll \mathrm{T}$, where T is the symbol period.

[22] Subjected to $|E_{ox}|^2|E_{oy}|^2 = \Re((E_{ox}E_{oy}^*)^2)$, the equality $S_0 = \sqrt{S_1^2 + S_2^2 + S_3^2}$ holds. This means that when the field projection along either X or Y is 0 or constant, or when both polarization components are identical, it is possible to detect S_{1-3} exclusively and thus reduce the number of of measured outputs to 3. Although in our case none of these requirements are met (note that all of them imply no more than 1 independent input), this receiver structure can be equally employed in all the coherent approaches introduced in Section 4.1, e.g. [227].

determinant influence on the robustness and simplicity of the DSP for polarization tracking and demultiplexing.

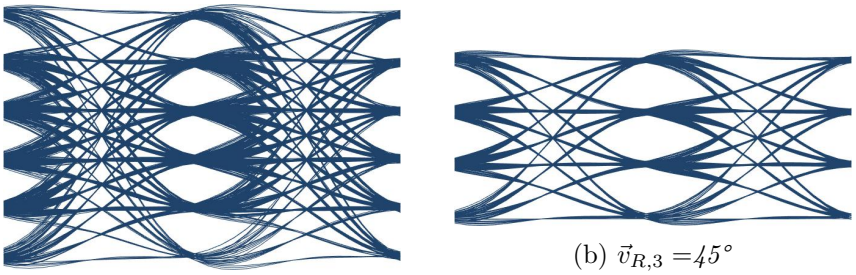
The $\vec{v}_{R,i}$ vectors must be selected to assure that the restoration of the Stokes parameters is possible; meaning that all potential received SOPs have a unique imprint and can be discerned. In our case, a standard Stokes analyzer was employed. In this structure, three of the polarizers are aligned to X or Y, *45°* or *135°* and RC or *left-circular (LC)* respectively. The remaining photodiode captures the instantaneous total field intensity (no polarizer). Given this setting, the Stokes parameters are calculated from the measured field intensities through the following linear transformation:

$$\begin{cases} S_0 = I_{TOTAL} \\ S_1 = 2(-1)^{K_1} I_{X/Y} - S_0, \text{ with } K_1 = \begin{cases} 1, & \text{if } Y \\ 0, & \text{if } X \end{cases} \\ S_2 = 2(-1)^{K_2} I_{45/135} - S_0, \text{ with } K_2 = \begin{cases} 1, & \text{if } 135^{\text{o}} \\ 0, & \text{if } 45^{\text{o}} \end{cases} \\ S_3 = 2(-1)^{K_3} I_{RC/LC} - S_0, \text{ with } K_3 = \begin{cases} 1, & \text{if } LC \\ 0, & \text{if } RC \end{cases} \end{cases} \tag{4.10}$$

where for the experimental demonstration, the polarizers are fixed to X, *135°* and RC for S_1, S_2 and S_3 respectively:

$$\mathbf{Rx} = \begin{pmatrix} \vec{v}_{R,1} \\ \vdots \\ \vec{v}_{R,M} \end{pmatrix} = \begin{matrix} - \\ X \\ 135^{\text{o}} \\ RC \end{matrix} \begin{matrix} \begin{matrix} X & \quad Y \end{matrix} \\ \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ \cos\frac{3\pi}{4} & \sin\frac{3\pi}{4} \\ \cos\frac{\pi}{4} & -i\cdot\sin\frac{\pi}{4} \end{pmatrix} \end{matrix} \tag{4.11}$$

It is clear from 4.10 that operating in the Stokes space implies a change to a coordinate system in which orthogonal SOPs ($X/Y \perp$ *45°/135°* $\perp RC/LC$) are not $X \perp Y$; or alternatively, multi-polarization signal design for Euclidean-distance-wise optimum symbol decision employing Jones basis (e.g. uniform angular distribution of linear SOPs [223, 224]) is different to that for optimum decision in the Stokes space. Besides being the essence of PolSK modulation, the latter fact evinces that orienting the transmitted SOPs along orthogonal Stokes co-



(a) $\vec{v}_{R,3}$ =*135°*

(b) $\vec{v}_{R,3}$ =*45°*

Figure 4.4: Simulated received eyediagrams in the absence of polarization mixing when four independent IM NRZ signals are transmitted according to 4.7. The eyediagrams are detected at the output of the photodetector when the preceding polarizer is aligned to *135°* (a) and *45°* (b). The inter-channel spacing is fixed at 100 GHz, the OSNR is 60 dB and the receiver bandwidth is 90% of the baud-rate. Both cases are presented in 1:1 proportion (horizontal axis: time; vertical axis: signal amplitude). The reduction in levels is apparent for the case where $\vec{v}_{T,j} \perp \vec{v}_{R,i}$ (6 levels against 4).

ordinates[23] avoids undesired intermediate vector transformations[24], as well as it improves the robustness of blind SOP-tracking algorithms based on Stokes-parameter manipulation. This motivates the particular set of transmitted Jones vectors $\vec{v}_{T,j}$ in 4.7, placing all the SOPs along the Stokes coordinates and maximizing the number of transmitted SOPs being mutually orthogonal in Stokes space. In addition, the number of orthogonal transmitted and detected Jones vectors ($\vec{v}_{T,j} \perp \vec{v}_{R,i}$) is intentionally maximized, thereby minimizing the order of the resulting multilevel signals under perfect alignment conditions. For instance, by orienting one of the $\vec{v}_{R,i}$ along *135°*, the coupling from the transmitted component along *45°* (90°degrees apart) is null (see Figure 4.4). The latter turns very convenient in the initial $\vec{v}_{T,j}$-$\vec{v}_{R,i}$ alignment and for system optimization.

[23]We refer to Stokes coordinates as those unit vectors describing the linear space spanned by the Stokes parameters. Being S_0 the aggregate scaling magnitude (see equation under footnote 17), Stokes coordinates are given by $\vec{S}_1^C = \begin{bmatrix}1 & 0 & 0\end{bmatrix}$, $\vec{S}_2^C = \begin{bmatrix}0 & 1 & 0\end{bmatrix}$, and $\vec{S}_3^C = \begin{bmatrix}0 & 0 & 1\end{bmatrix}$; whose respective Jones vectors are $\vec{v}_{S_1^C} = \begin{bmatrix}1 & 0\end{bmatrix}$, $\vec{v}_{S_2^C} = \begin{bmatrix}\cos\frac{\pi}{4} & \sin\frac{\pi}{4}\end{bmatrix}$ and $\vec{v}_{S_3^C} = \begin{bmatrix}\cos\frac{\pi}{4} & -i\cdot\cos\frac{\pi}{4}\end{bmatrix}$.

[24]Representation of whatever $\vec{v}_{T,j}$ according to the orthogonal Stokes basis $\vec{v}_{R,i}$ in which polarization tracking and rotation compensation is done.

4.2.5 Digital receiver - stages and practical considerations

The DSP chain described in this subsection corresponds to one of the possible solutions for satisfactory polarization tracking and data demodulation (considering FEC codes) employing the proposed fourth-order polarization-multiplexed IM/DD system. We subdivide the elaboration in three main blocks: (i) general front-end correction, (ii) polarization tracking plus channel rotation compensation and (iii) demultiplexing. The full DSP schema is included in Appendix D for the readers' convenience.

4.2.5.1 Front-end correction and signal conditioning

The purpose of the first set of algorithms within this block is to prepare the digitized photocurrents in order to maximize the accuracy of the subsequent calculation of Stokes parameters (see 4.10) for best-possible SOP tracking. The comprehended steps are:

Resampling and Filtering: The four sequences are resampled to an integer number of samples per symbol (2 in our case). Out-of-band noise and the remaining undesired spectral components caused by signal-signal beating in the photodetector are attenuated for SNR maximization.

Time skew : Skews originating both at the transmitter and receiver sides have different though critical impact on the system's functioning. The skew at the transmitter cannot be easily corrected by DSP at the receiver, and it requires precise symbol synchronization between the independent signal generators up to the fiber input. It determines the quality of the received multilevel signals, affecting both the cleanliness of the levels and the sampling-point deviation tolerance. Note that the relative delay between carriers caused by chromatic dispersion appears as transmitter skew, evincing that every frequency grid requires specific compensation.

The skew at the receiver does not alter the multilevel signals' quality as described above, but their respective time alignment before digitization. Because the calculation of all the Stokes parameters depends on the measured instantaneous total field intensity (S_0), misalignments of sufficient magnitude between symbol periods can severely affect the performance. For instance, a relative skew of S_0 with respect to S_{1-3} of $>30\%$ the symbol time prevents successful demodulation owed to the errors induced by

the SOP tracking algorithm[25]. In addition, because the total propagated power is not affected by channel-induced polarization mixing, S_0 becomes a robust reference for common optimum sampling selection irrespective of the severity of the distortion. This underscores the need for accurate symbol synchronization across all the four sequences, where preceding frame/coarse synchronization is understood.

Scaling : The amplitude ratios between the measured photocurrents are crucial for the precise determination the polarization-mixing coefficients. Therefore, the individual normalization of the captured photocurrents is a mistake that leads to distorted Stokes parameters when the intensity waveforms are combined (4.10). It follows that differences in the responsivities of the four photodiodes must be equalized *a priori*, in order to avoid incorrect impairment compensation and demultiplexing.

DC offset : Different photodiodes may also show disparate output DC components for the same input optical power. Even with equalized responsivities, uncompensated DC values may accumulate when calculating the Stokes parameters, then producing biased waveforms that modify the symmetry of the decision boundaries. Unless addressed, this effect compromises the robustness of the algorithm for SOP tracking and demultiplexing. Note that the response of the photodiodes may also vary with the input power level. In such case, scaling factors and DC compensation values should be adjusted according to the received optical power and/or power-dependent adjustment of the slicers' thresholds should be implemented for improved decisions.

After the first set of corrections on the digitized photocurrents, the transformation of the intensities to Stokes parameters is done as indicated in 4.10. From this point on, the DSP operates on the Stokes parameters:

Timing recovery and sampling point : As previously mentioned, the invulnerability of S_0 to polarization mixing enables its use as robust reference for common clock recovery and optimum sampling point estimation. This is conditional upon precise skew compensation and negligible deviation between the individual sampling clocks across the four branches. In the experimental demonstration, standard Gardner error detector with

[25]The number is estimated from simulated noiseless B2B configuration with 100-GHz inter-channel spacing and ideal signal-signal beating removal.

proportional-integrator loop is employed for clock recovery followed by maximum-variance decimation [87].

4.2.5.2 Polarization mixing compensation

In this stage, the time-varying Mueller matrix of the channel $\mathbf{M_{dist}}$ is blindly estimated ($\mathbf{H_{dist}}$ in Jones calculus[26]) for the subsequent compensation of possible polarization mixing, thus enabling effective SOP demultiplexing and demodulation. Because channel-induced polarization mixing reflects as a 3D rotation (see footnote 21) of the sub-space defined by S_{1-3} (graphically represented by the Poincaré sphere), the channel's Mueller matrix reduces to the standard rotator form:

$$\mathbf{M_{dist}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & S_{1,1} & S_{1,2} & S_{1,3} \\ 0 & S_{2,1} & S_{2,2} & S_{2,3} \\ 0 & S_{3,1} & S_{3,2} & S_{3,3} \end{pmatrix}, \tag{4.12}$$

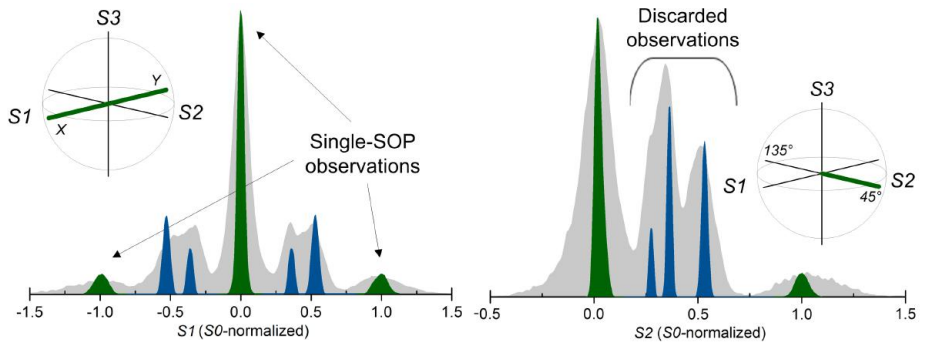
where $S_{i,j}$ for $i, j \in \mathbb{Z}$ are real-valued coefficients which quantize the normalized coupling factors between the Stokes parameters i and j.

The next algorithm is inspired by the work in [195], where a blind method for compensating polarization-mixing on a single Stokes coordinate was presented (2^{nd}, 3^{rd} or 4^{th} row in 4.12, see footnote 23); thereby enabling PolMux of up to two orthogonal Jones vectors undergoing very fast polarization rotation. The main DSP-novelty introduced in this work resides in the capability to measure the rotations on all three Stokes coordinates, that is, tracking $\mathbf{M_{dist}}$ entirely. This extension maintains the simplicity and blindness of the original algorithm while making possible polarization tracking and de/multiplexing beyond two SOPs which, additionally, are not restricted to be orthogonal.

In the following, the details of the tracking process are elaborated. This is supported with Figure 4.5, where simulated (color) histograms of the normalized Stokes parameters S_1 (a) and S_2 (b) are shown for perfect

[26] For Jones-to-Mueller conversion: $\mathbf{M} = \mathbf{A}(\mathbf{J} \otimes \mathbf{J}^*)\mathbf{A}^{-1}$, where $\mathbf{M}$ and $\mathbf{J}$ are the Mueller and Jones matrices respectively and $\mathbf{A}$ is given by:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{pmatrix}$$



(a) S_1 (b) S_2

Figure 4.5: Histogram of the Stokes parameters S_1 (a) and S_2 (b) normalized by S_0 for perfect transmitter-receiver polarization alignment in the polarization-multiplexed IM/DD data transmission system described by 4.7 and 4.11 employing NRZ modulation. In color, simulated data for 40-dB SNR B2B configuration. In grey, experimental results for 4×27 GBd at 6 dBm optical power into the receiver. Insets: Poincaré spheres. The Stokes parameters of both (a) and (b) are highlighted.

polarization-mixing compensation. Experimental results (grey) are presented for comparison. Note that the asymmetry around zero of Figure 4.5b is due to the fact that $45°$ is the only $\vec{v}_{T,j}$ explicitly lying along $\vec{S}_2^C$. This contrasts with the projection on $\vec{S}_1^C$, since both X and Y are transmitted. In this regard, S_2 and S_3 show similar histograms.

SOP tracking : First, the trivial zero-power cases are detected and sliced by thresholding on S_0's lowest level. The samples below threshold can already be demapped to the corresponding bit tuple (e.g. [0 0 0 0] with each column associated with one of the four independent data streams), and the respective instantaneous samples are removed from the four Stokes-parameter sequences. This avoids mathematical indeterminations in the following normalization by S_0, and slightly reduces the processing load by discarding trivial cases to track on. For its part, normalizing by S_0 compresses the Stokes space to the unit-radius sphere irrespective of the modulation format, the transmitted power or the polarization-dependent distortions. Since the characteristic matrix of the channel is assumed unitary, common S_0-normalization becomes convenient to simplify the channel-estimation algorithm by removing the influence of the irrelevant magnitude

variable.

Because the number of possible magnitudes and orientations (Euler angles) of the aSOP is finite and deterministic given the modulation format per SOP, the representation of the output combinations in the Stokes space shows well-determined locations in the absence of distortion or after polarization-rotation compensation. Therefore, the deviations with respect to such pre-defined positions serve as quantifiable figures of the experienced polarization rotation. This is the observation governing the algorithm for estimating the channel's characteristic matrix ($\mathbf{M_{dist}}$).

As first step, the projection of all the aSOP combinations upon the Stokes coordinates in the absence of polarization mixing must be identified and studied. In this regard, we subdivide the cases into two groups:

(A) *Impinging aSOP aligned to one $\vec{S}^{C}_{1-3}$ and conformed by 1 non-zero SOP* : all the power is projected onto the respective Stokes coordinate (the magnitude of one Stokes parameter equals 1 and the others' equals 0), thus resulting in the maximization of the Euclidean distance between the magnitude levels. The latter is visible in the histograms in Figure 4.5, where the discussed (signed) cases are displayed in green color. For example, the observations around 1 in the S_1 histogram correspond to those instants when only $\vec{v}_{T,1}$ was excited (e.g. [1 0 0 0] input symbol-tuple), in turn making the received aSOP be oriented along X. Those very same observations contribute to the point accumulation around zero in the S_2 histogram. In these instances aSOP=SOP.

(B) *Impinging aSOP conformed by any number >1 of non-zero SOPs* : at these time instants the magnitude of the individual Stokes parameters describing the received aSOP vector will be <1 (blue observations in Figure 4.5). Unlike in the previous group, these magnitude levels vary according to the relative power ratios of the individual SOPs, hence requiring dedicated detection-thresholds for each modulation format. Furthermore, the higher proximity between the projections as compared to the previous group, increases the overlapping region among the noisy clusters (see Figure 4.5); hence hindering their detection and thus the robustness of SOP tracking.

For improved robustness in blind channel estimation, and for enabling the generalization of the SOP tracking algorithm to other modulation formats, the observations belonging to group B are disregarded in this work.

As a result, the channel matrix $\mathbf{M_{dist}}$ can only be estimated conditional on $\vec{v}_{T,j}$ defining a collection of SOPs with at least one oriented along each Stokes coordinate. In the proposed 4-SOP system, $\vec{v}_{T,1} \perp \vec{v}_{T,3}$, $\vec{v}_{T,2}$ and $\vec{v}_{T,4}$ are aligned with each of the three Stokes coordinates (see 4.7 and footnote 23). It is important to emphasize that because only group-A observations are used for error signal estimation, polarization tracking is not totally continuous. The probability of occurrence of these specific aSOPs will determine the tracking accuracy around the respective coordinate for a given observation time interval (more samples for filter adaptation); whose duration is dictated by a compromise between the baud-rate, the polarization rotation speed, and the tolerable penalty caused by the polarization mixing. Nonetheless, given uniform probability distribution of the input data streams, the speed at which the three necessary kind of group-A observations occur at the considered baud-rates is much faster than the speed of polarization changes in regular fiber communications (below microsecond scale [225]). Given the proposed 4-SOP system with $\vec{v}_{T,j}$ defined by 4.7, the probability of transmitting a group-A aSOP along each axis:

$$\begin{cases} P_{\vec{S}_1^C} = \frac{2}{16} \ (X \text{ and } Y) \\ P_{\vec{S}_2^C} = \frac{1}{16} \ (45°) \\ P_{\vec{S}_3^C} = \frac{1}{16} \ (LC) \end{cases}$$

where the probability of receiving all three is:

$$P_{TOTAL} = P_{\vec{S}_1^C} \cdot P_{\vec{S}_2^C} \cdot P_{\vec{S}_3^C} = \frac{1}{2048} \tag{4.13}$$

The conclusion is now immediate: after locking, the entire coordinate system can be updated every 2048 symbols. At a transmission speed of 25 GBd, the updating periodicity ≈0.082 μs. Accordingly, the updating assiduity will be deemed quasi-continuous, as it is in the order of 100 times faster than required.

As previously stated in Section 4.2.4, polarization rotation appears as polarization mixing when the signal is detected with a static polarization-diversity receiver. That being so, the total received power spreads among the three Stokes coordinates even when the transmitted aSOP is perfectly aligned with one of them. This effect reflects as a unitary 3D rotation of the real vector space determined by $\vec{S}_{1-3}^C$. Particularizing to observations belonging to group A, the deviation from the unit magnitude along the

corresponding Stokes coordinate $\vec{S}^{C}_{1-3}$ at every instant (together with the leakage on the others and their relative ratios) represents a time-varying quantification of the observed polarization rotation around that particular axis. This information (error signal), shall be employed to update the coordinate system accordingly so that, under these new reference axes, the received Stokes vector appears uncoupled. Therefore, polarization mixing compensation starts with the blind estimation of the signed error signal for each reference vector, whose mathematical expression is given by (after [195]):

$$\vec{e}_{1-3}[n] = \begin{cases} \text{sign}(\vec{S}^{E}_{1-3}[n]) \frac{\vec{S}[n]}{S_0[n]} - \vec{r}_{1-3}[n], & \text{if } |\vec{S}^{E}_{1-3}[n]| > \text{K} \\ 0, & \text{otherwise} \end{cases} \tag{4.14}$$

with

$$\vec{S}^{E}_{1-3}[n] = \left(\frac{\vec{S}[n]}{S_0[n]} \right) \vec{r}_{1-3}[n] \tag{4.15}$$

where $\vec{S}[n]$ is the received Stokes vector, $S_0[n]$ is the total intensity, $\vec{r}_{1-3}[n]$ are the unitary reference vector (progressively updated according to the estimated rotation) and $\vec{S}^{E}_{1-3}[n]$ are the projections of the normalized received Stokes vector along $\vec{r}_{1-3}[n]$. For its part, K provides static threshold for selecting group-A observations exclusively which, employing Figure 4.5 as example, $|\text{K}| \approx 0.8$. It is likewise apparent from Figure 4.5 that, although group-B observations are disregarded for updating the error signal, they may hinder the detection of group-A instants under the influence of noise (notice the overlap between the equivalent green and blue regions in the experimental data). This emphasizes the importance of accurate threshold calculation, as well as it evinces its dependence on $\vec{v}_{T,j}$ and the modulation format [197].

The three ensuing signed error signal are then included in respective normalized least mean square (NLMS) adaptive filter calculations targeting the adaptation of $\vec{r}_{1-3}$ to mimic the time-varying rotation undergone by $\vec{S}$ with respect to the Stokes coordinates $\vec{S}^{C}_{1-3}$:

$$\vec{r}_{1-3}[n+1] = \frac{\vec{r}_{1-3}[n] + \mu \vec{e}_{1-3}[n]}{\|\vec{r}_{1-3}[n] + \mu \vec{e}_{1-3}[n]\|} \tag{4.16}$$

where μ stands for the convergence step size and $\| \bullet \|$ is the norm operator.

These calculations are executed for each of the three coordinate vectors in series, where the alternation is triggered by a zero-power exception indicating that any other orthogonal Stokes coordinate has been transmitted instead (green-colored observations around 0 in Figure 4.5). These serial operation is embedded in an iterative loop which breaks when the error signal is converged (for which a user-defined threshold must be pre-established) or when the group-A observations within the captured window are exhausted. It is worth remarking that the algorithm described here is a tracking tool and, as such, assumes prior initialization (e.g. via training sequences [199, 212, 228]). Assuming initialization, the algorithm locks after convergence, then enabling simple, quasi-continuous, and blind characterization of the polarization rotation for all the Stokes coordinates. The process ends by building 4.12.

Channel distortion compensation : In the SOP-tracking process, the three-dimensional rotation of the Stokes space is characterized and quantified. Equivalently, a polarization-mixing-free coordinate system is obtained. Assuming the inability to change $\vec{v}_{T,j}$ and/or $\vec{v}_{R,i}$ dynamically, we realize digital derotation of the received Stokes vector $\vec{S}[n]$. Given that rotation does not break the orthogonality among the coordinate vectors, the matrix inverse is readily calculated through transposition:

$$\mathbf{M}_{\mathbf{dist}}^{-\mathbf{1}} = \mathbf{M}_{\mathbf{dist}}^{\top} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & S_{11} & S_{21} & S_{31} \\ 0 & S_{12} & S_{22} & S_{32} \\ 0 & S_{13} & S_{23} & S_{33} \end{pmatrix} \tag{4.17}$$

Afterwards, the matrix is applied on the received Stokes sequences. From this step on, transmitter and receiver are assumed perfectly aligned ($\mathbf{H}_{\mathbf{dist}} = I_2$, $\mathbf{M}_{\mathbf{dist}} = I_4$).

4.2.5.3 Demultiplexing

The third and last block deals with static polarization demultiplexing. This stage comprises Stokes-to-intensity transformation and *demapping matrix* application. Because both mentioned transformation matrices are known a priori, no real-time inversion is involved. The two aforementioned processes can be done separately, or they can be collapsed into a single pre-defined and static matrix multiplication. In both cases, the four SOPs are decoupled and the estimated transmitted intensity waveforms are retrieved. The

inputs are the Stokes parameters at the output of the channel compensation step ($\vec{S}^{E}_{1-3}[n]$) irrespective of the approach:

- *Two steps*

  First, the input Stokes parameters are transformed into intensities through 4.10. After that, the *demapping matrix* is applied.

- *One step*

  The matrix $\vec{v}_{T,j}$ is converted to Stokes parameters, and its inverse is directly applied on $\vec{S}^{E}_{1-3}[n]$ without intermediate conversion to intensity waveforms. For the particular configuration 4.7, the demultiplexing matrix is given by:

$$\mathbf{M_{demux}} = \begin{pmatrix} 0.5 & 0.5 & -0.5 & -0.5 \\ 0 & 0 & 1 & 0 \\ 0.5 & -0.5 & -0.5 & -0.5 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.18}$$

The DSP is ended with a slicer and error counting before FEC. Note that other features/technologies can be added on top of the presented stages (e.g. nonlinear equalization for quadratic noise compensation or support for higher-order modulation formats) or in place (e.g. multi-dimensional clustering algorithms for joint Stokes vectors tracking). However, the proof-of-concept experimental demonstration was realized with the strictly necessary modules for simplicity and worst-case evaluation.

4.3 Experimental demonstration

Critically driven by cost and power consumption, client-side optical transceiver modules are designed upon stringent requirements on the maximum average launch power. For instance, Ethernet task forces IEEE P802.3ba (40 Gbps and 100 Gbps) and IEEE P802.3bs (400 Gbps) specify quaternary WDM configurations with $\leq$4 dBm per lane [177,229]. Together with the various power-budget margins accounting for manifold purposes, those restrictions establish an upper bound for the receiver sensitivity beyond which communication is not sustainable. Therefore, characterizing the receiver sensitivity is essential for assessing the practical interest of data-communications systems intended for application in closed short-reach

networks. This is the purpose of the following section, where the experimental validation of the 4-SOP IM/DD data transmission system described in Section 4.2 is elaborated.

Targeting 100 Gbps net bit-rate, 4×27 GBd and 4×32 GBd NRZ configurations are considered along with 7% and 20% overhead FEC codes respectively. The performance of both B2B configuration and 2-km SSMF transmission is evaluated in terms of BER versus received optical power, whereas the robustness of the SOP-tracking algorithm is analyzed in terms of time-averaged BER under time-varying polarization changes far more demanding than expected in the considered scenarios [225, 230]. Throughout the demonstration, only the strictly necessary DSP blocks are employed (see Section 4.2.5).

4.3.1 Testbed

Figure 4.6 shows the schematic of the experimental setup. At the transmitter side, four C-band distributed feedback lasers ($\text{DFB}_{\text{A-D}}$) with 100-GHz spaced center frequencies (193.2-5 THz) are used as light sources. Frequency spacing is used to ensure incoherent power addition after their beating in the receiver's photodiodes[27]. The outputs of the distributed feedbacks (DFBs) are externally modulated with four integrated Mach-Zehnder modulators (MZMs) with ≈30-GHz eletro-optical bandwidth. Note that the system employs IM, opening for the use of other external-modulation devices or direct-modulation techniques. The MZMs are electrically driven by two pulse pattern generator (PPG) modules, delivering a total of four independent 27 GBd (108 Gbps gross rate) or 32 GBd (128 Gbps gross rate) NRZ data streams with ≈3 Vpp. Standard PRBSs of length $2^{15}-1$ are used, and fine symbol synchronization is attained with electrical delay lines. No DSP is employed at the transmitter. Lastly, the SOP of each modulated carrier is aligned along the respective $\vec{v}_{T,j}$ (i.e. X,

[27]For a given bandwidth per transmitted carrier, the laser frequency spacing will determine the degree of interference from the higher-frequency beatings to the baseband component. Insufficient separation will make these higher-frequency beatings fall within the photodiodes' electrical bandwidth, thus partially overlapping the baseband signal with the consequent distortion (the modulation bandwidth is assumed greater or equal than the photodiodes' electrical bandwidth).

Interference between non-orthogonal Jones vectors occurs when the lasers' frequency spacing ranges from DC to twice the electrical bandwidth of the photodiode. In non-severe cases, digital filtering can minimize the impact of the distortion. Severe cases may prevent correct signal demodulation, hence requiring the redesign of the frequency grid as first approach.

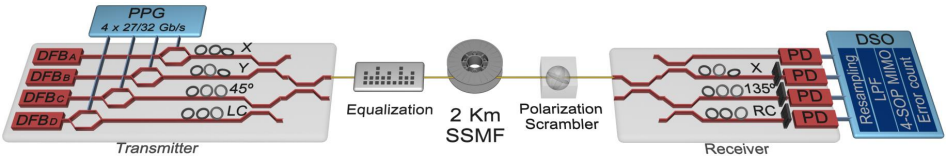


Figure 4.6: Experimental setup of the considered 4-SOP data transmission system.

Y, *45°* and *LC*, see 4.7) with polarization controllers, and combined with one 4×1 optical coupler.

The output from the transmitter is passed through a power equalization stage before fiber transmission. This stage comprises an EDFA operated in saturation regime (≈5 dB of noise figure) plus a variable optical attenuator (VOA). This structure allows for compensation of the disparate insertion losses among the MZMs in our transmitter, in no case providing input-output optical gain. After equalization, the signal is launched into 2-km SSMF for transmission. One polarization scrambler is installed at the end of the link for the evaluation of the polarization tracking algorithm.

At the receiver side, the signal is detected with a standard *Stokes analyzer* (see $\vec{v}_{R,i}$ in 4.11) with ≈30-GHz electrical bandwidth. The top branch in Figure 4.6 measures the instantaneous field intensity (S_0), the other three include polarization controllers (PolCs) and polarization beam splitters (PBSs) for providing the required polarization diversity (Section 4.2.4). After photodetection, the four electrical signals are digitized and stored for offline processing in a DSO with 25-GHz bandwidth input ports and 80-GS/s sampling rate. The DSP at the receiver consists of front-end correction, Stokes-based compensation of the polarization mixing plus polarization demultiplexing, and error counting (see Section 4.2.5 for detailed explanation).

4.3.2 Results

The discussion on the results is subdivided into (i) receiver sensitivity analysis for both B2B and transmission configuration, (ii) polarization tracking performance, and (iii) comments on power consumption and processing complexity.

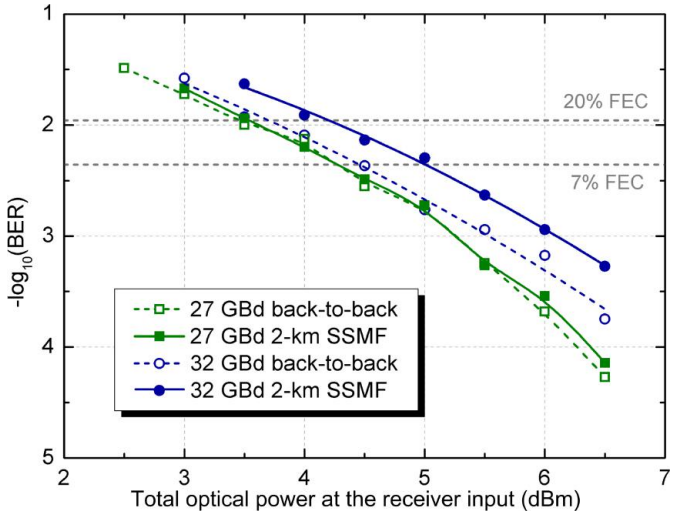


Figure 4.7: BER (≈2.9 million bits per point) versus total and per-channel optical power at the input of the receiver for B2B and 2-km SSMF transmission.

4.3.2.1 Receiver sensitivity

Figure 4.7 shows the BER performance of the system versus optical power at the input of the receiver for perfect polarization rotation compensation (polarization tracking algorithm always operational). Measured curves for both B2B and transmission after 2-km SSMF are presented for the two cases of study, i.e. 4×27 GBd (108 Gbps) and 4×32 GBd (128 Gbps) gross rates. The pre-FEC BER thresholds for post-FEC BER of 10^{-15} are included for hard-decision 7%- and 20%-overhead BCH codes ($4.4 \cdot 10^{-3}$ [231] and $1.1 \cdot 10^{-2}$ [232] respectively). Aiming 100 Gbps net bit-rate, lower baud-rate is associated with lower code-rate in this work; yielding net bit-rates of ≈100.93 Gbps for the 4×27 GBd case, and ≈106.66 Gbps 4×32 GBd. Besides overhead, processing complexity and latency, and de/coding power consumption are other important FEC-dependent aspects that should be considered for optimum transceiver design. These are, however, hardly relevant from the receiver sensitivity perspective.

Below-FEC threshold BER is attained for both 27 GBd and 32 GBd irrespective of the transmission configuration. For 27 GBd, insignificant transmission penalty is observed, what makes the overall receiver sensitivity ≈4.4 dBm (≈-1.6 dBm per wavelength). For its part, 32 GBd B2B

exhibits similar performance to 27 GBd below ≈5 dBm, showing a receiver sensitivity of ≈3.7 dBm (≈-2.3 dBm per wavelength). The reason for the divergence beyond ≈5 dBm is that, as SNR increases with average launch power (nonlinearities are negligible due to the short propagation), the ISI originating from signal over-filtering starts dominating the error rate; thus causing a slower BER decay rate in the high baud-rate signal. Note that zero-tap filters are employed for channel equalization (see 4.17), thus intra-channel ISI is not compensated in this experiment. The penalty after transmission is ≈0.5 dB, what worsens the pre-FEC sensitivity to a range comparable to that of 27 GBd at the 7%-overhead FEC threshold. This leads to similar power budget margin between both baud-rate cases, which could be argued to favor 27 GBd over 32 GBd in scenarios targeting 100 Gbps (despite the 2 Gbps lower transmission rate for 27 GBd) based on the lower processing complexity of the higher code-rate, and the lower requirements in electrical bandwidth.

Benchmarking : For putting into perspective the above-described sensitivity results, we resort to the investigation in [233]; where we compare the 4×32-GBd case against analogous transmission systems employing parallel optics (one fiber per carrier) and WDM technology (filter-based wavelength demultiplexing). The modulation format, the inter-channel spacing, and the considered FEC are therefore kept in NRZ, 100 GHz and hard-decision 20%-overhead BCH respectively. As far as DSP is concerned, only static front-end corrections and timing recovery are preserved for parallel optics and WDM configurations as they do not feature polarization diversity (see Section 4.2.5). Figure 4.8 presents the sensitivity results after [233] in BER versus total optical power (four channels) at the input of the receiver[28]. For brevity and because the difference between the curves of WDM and parallel optics reduces to a constant ≈1-dB sensitivity degradation owed to optical filtering, we will only compare the results of the 4-SOP system with those of parallel optics.

The experimental data for B2B configuration shows ≈14.9 dB sensitivity penalty at the pre-FEC BER threshold. This number results from (i) the different power loss from the input of the receiver to the input of the photodiodes between both cases[29] and (ii) the demultiplexing process (see

[28]For the case of parallel optics, the individual power of each carrier is individually measured and added a posteriori due to the physical separation of the channels.

[29]The insertion losses owed to the extra filter in the WDM case are already corrected

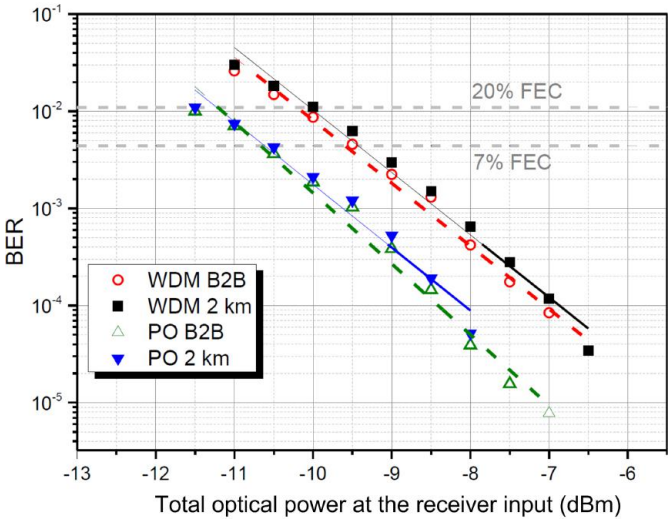


Figure 4.8: BER versus received optical power at the input of the photodiode for parallel optics and WDM technology considering both B2B and 2-km SSMF transmission [233].

Section 4.2.5.3), which enhances the noise per channel by aggregating different scaled combinations of the independent AWGN components in the receiver branches:

- *Insertion loss:* referencing to parallel optics, the 4-SOP system presents one extra 1×4 optical power splitter, plus PolCs and PBS for the three bottom branches in Figure 4.6; all in all yielding: 7.1 dB (S_0), 8.8 dB (X), 9.1 dB ($135°$) and 8.9 dB (RC) extra losses. Because the field intensity along Y is reconstructed from S_0 (low loss) and S_1 (high loss) through $I_Y = \frac{S_0 - S_1}{2}$, it exhibits slightly better average performance than the explicitly detected SOPs (high loss). Nevertheless, for simplicity and taking into account that the BER curves in Figure 4.7 result from the average of all four channels, the sensitivity degradation associated to this contribution is approximated to 9 dB for all branches.

- *Demultiplexing:* so as to quantify noise enhancement, we refer to 4.2 and 4.9, and assume independent zero-mean and same-variance AWGN across the four branches $N_j(0, \sigma^2)$. Note that the unitary

for in Figure 4.8

rotator matrix employed for channel distortion compensation ($\mathbf{M_{dist}^{-1}}$) does not have any influence concerning noise enhancement, hence it is made equal to I_4 for simplicity:

$$\mathbf{W} = \begin{pmatrix} 1.0 & -1.0 & 0.0 & 1.0 \\ -1.0 & 2.0 & -1.0 & 0.0 \\ 0.0 & -1.0 & 1.0 & 1.0 \\ 1.0 & 0.0 & 1.0 & -2.0 \end{pmatrix} \begin{pmatrix} N_X(0,\sigma^2) \\ N_{45^\circ}(0,\sigma^2) \\ N_Y(0,\sigma^2) \\ N_{LC}(0,\sigma^2) \end{pmatrix} = \begin{pmatrix} N_X(0,3\sigma^2) \\ N_{45^\circ}(0,6\sigma^2) \\ N_Y(0,3\sigma^2) \\ N_{LC}(0,6\sigma^2) \end{pmatrix} \tag{4.19}$$

Four new and independent noise processes with apparent increased variance are defined for each intensity signal. The values correspond to a sensitivity of ≈4.7 dB for X and Y and ≈7.8 dB for $45°$ and LC. Again considering that the BER curves in Figure 4.7 result from the average of all four channels, ≈6-dB sensitivity penalty shall be expected.

Summing both average contributions, ≈15 dB sensitivity penalty is obtained, precisely agreeing with the outcome from the experimental results and confirming that, in 4-SOP systems, the receiver sensitivity suffers notorious and unavoidable degradation even under ideal polarization rotation compensation (or absence of channel distortions). This extra optical power requirement greatly contributes to the overall power consumption through lasers driving, in turn representing up to ≈40% of the total [233].

For its part, the difference between the transmission penalty for parallel optics (negligible) and the 4-SOP system (0.5 dB) reports about the inaccuracies caused by CD on the channel estimation process. This gap to B2B can be reduced via multi-tap signal equalization, which along with pulse-shaping techniques, can effectively reduce the impact of CD and any other form of linear ISI (e.g. signal over-filtering at the electro-optical and opto-eletrical interfaces) at the expense of processing complexity. In addition, improvements on the insertion loss of the optical components (e.g. power splitter/combiner or modulators) will increase the power-budget margin, thus enhancing the protection against aging, imperfect connectors and splicings, and fiber attenuation. In turn, this would motivate operating in the O-band as a means of canceling out the influence of CD and then reducing the DSP complexity.

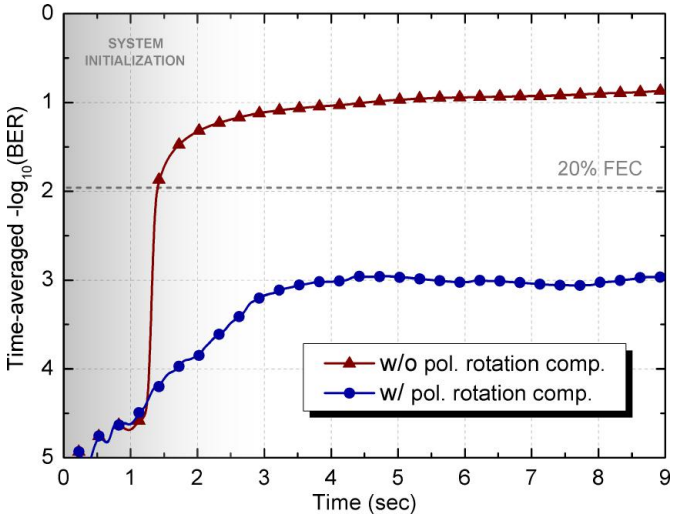


Figure 4.9: Time-averaged BER versus time in B2B configuration for 4-SOP 32-GBd transmission with (blue) and without (red) polarization-rotation compensation algorithm.

4.3.2.2 Polarization tracking

The experimental assessment of the polarization tracking algorithm is done with the help of an inline polarization scrambler. The fastest rotation rate the scrambler can achieve is on the order of radians per second (≈18 rad/s mode and ≈200 rad/s after [225]), which is much slower than what can be detected within the maximum observation window of the employed DSO. This means that one single frame could not capture such polarization changes even using the maximum-available storage memory. To overcome this practical impediment, short traces are periodically stored, where the time-period determines the rotation magnitude to be tracked given fixed angular speed and how often such estimation is done. After calculating the rotation matrix, it remains constant for the entire period until re-update. This routine allows for testing the software for diverse rotation magnitudes (i.e. time-periods), with a limit determined by the minimum necessary number of symbols for successful tracking ($\leq$2048 symbols in our case, see *SOP tracking* in Section 4.2.5.2) and the baudrate. Note that this approach is equivalent to fixing the time-period and increasing the angular rotation speed.

In our experimental characterization, the angular speed is fixed at max-

imum, which is around one order of magnitude faster than expected in buried-fiber scenarios and four times faster than aerial routing [230]. A total of 120 frames of 16000 bits each were processed every 75 ms (total of 9 seconds). The measurement process starts with and initialization period comprehending ≈1 second during which the polarization scrambler remains idle and $\vec{v}_{T,j}$ are kept perfectly aligned to $\vec{v}_{R,i}$. Figure 4.9 presents the time-averaged BER over time with and without channel estimation (the same data sequence is used) for the 4-SOP×32-GBd configuration at 25.5 dB SNR subjected to the above-described polarization rotation. This case is deemed representative of poor performance case.

After the initialization interval, the received SOP progressively deviates from ideal alignment. In the case without polarization tracking, the instantaneous BER quickly converges to ≈0.5 (with the exception of those coincidental moments in which the received SOP is perfectly aligned with the receiver). Consequently, the time-averaged BER shows an apparent monotonically increasing tendency. However, when the same data is processed with polarization tracking the performance changes considerably. After initialization, the instantaneous BER starts oscillating within a maximum of $\approx 8{\cdot}10^{-3}$ and a minimum of 0 (see[30]). This instability suggests using shorter updating periods for the channel matrix (<16000 symbols) and dynamic (impairment dependent) threshold control (see K in *SOP tracking* in Section 4.2.5). The resulting time-averaged BER shows stable around $\approx 10^{-3}$, denoting successful demodulation after 20%-overhead FEC processing[31]. The presented curves prove (i) the need for rotation compensation to accomplish correct signal recovery, and (ii) successful post-FEC demodulation with the algorithm proposed in Section 4.2.5 under conditions much more demanding than can be expected short-reach transmission links [230].

4.3.2.3 Power consumption and processing complexity

Despite ≈40% of the total power consumption comes from lasers driving in the proposed 4-SOP system, there is a considerable remaining fraction that cannot be neglected. Based on the results collected in [233], it is precisely the contribution from this remaining part what makes the major difference, yielding around ≈37% higher estimated power consumption

[30] Because 16000 symbols are employed for each measurement point in Figure 4.9, the reliable BER is $\approx 6{\cdot}10^{-4}$. Zero counted BER must be interpreted cautiously.

[31] This can be assured because the worst-case instantaneous BER does never exceed the threshold for 20%-overhead FEC. Not being the case, successful demodulation is conditional upon error spreading via bit-interleaving or similar techniques.

per channel when referenced to parallel optics. This is because in 4-SOP, four-fold more sequences need to be digitized through respective ADCs for recovering a single channel. Note that the number of ADCs is identical in terms of hardware; that is, if the channels are employed to carry in parallel mode a single higher-speed serial data stream belonging to the same user (in contrast to one user per carrier), the additional power consumption vanishes.

Interestingly, the polarization-dependent DSP only constitutes 2.47% additional power consumption, confirming the good effectiveness-to-simplicity ratio of the proposed algorithm. Translating to complexity, both the extra number of real multiplications and additions keep below 1%. Enhanced equalization, or the adaptation to AMF will surely increase the consumption which, nevertheless, is not expected to exceed 5-7%.

In order for polarization-based wavelength demultiplexing to be worth consideration, all additional contributions to the power consumption must not exceed the one associated to standard WDM-demultiplexing apparatus. Although this balance is not included in this report, the reader is referred to [233] for a detailed breakdown of the power consumption and the processing complexity.

4.4 Summary

Fundamentally driven by cost and power consumption, throughput scaling in client-side optics is sustained by, first, optimizing the resources for serial rate maximization; and subsequently, incorporating new dimensions when possible and needed. This routine has based the upgrade and development of transmission standards across all currently considered reaches and line-rates, giving rise to hybrid serial/parallel IM/DD transceiver modules featuring both sustainable power consumption and very compact form factors.

Primarily owed to the simplicity of the demultiplexing process and the cost-effectiveness of the associated equipment, WDM and parallel optics have been the preferred approaches for throughput aggregation. Albeit these technologies have sufficed to overcome electrical-bandwidth limitations as yet, the efficiency of their scalability saturates as the terabit traffic range is approached; excessively increasing the number of parallel lanes and leading to yield degradation, challenging packaging, and aggravating the requirements on lasers' frequency stability and thermal control. In consequence, the incorporation of techniques that allow for increasing the

aggregate throughput without employing additional lasers or fiber pairs becomes of great interest. This has recently put the research focus on quadrature- and/or polarization-multiplexing, from which the latter constitutes the area covered in this investigation.

This investigation

In this chapter, we have described in detail a PolMux IM/DD technique enabling transmission and demultiplexing of up to four distinct SOPs in fiber-optic communication systems, and with the potential to be used alongside or in place of existing technologies (e.g. WDM or AMF). In the first block the essentials of this technique are discussed, covering from general mathematical description of N-fold PolMux, to requirements on the transmitter plus receiver structures. The study intuitively proves that the maximum number of simultaneously transmittable SOPs equals four, and elaborates on their optimum orientation with respect to the polarization-diversity receiver for (i) facilitating system optimization, (ii) increasing the robustness and simplicity of the polarization tracking algorithm operating on Stokes parameters and (iii) improving the overall error rate by minimizing noise enhancement. Additionally, one potential DSP chain for blind polarization-rotation tracking and signal demultiplexing based on Stokes parameter manipulation is thoroughly explained, allowing quasi-continuous (microsecond-periodic) channel estimation at low computational complexity.

Besides the traditional $X \perp Y$ dual-polarization approach for SE doubling, non-orthogonal ($\geq$3-SOP) PolMux configurations are suggested as alternative demultiplexing method in other multiplexing technologies; then permitting the interchange of the respective de/multiplexing apparatus. In this regard, we propose and experimentally demonstrate the combination of 4-SOP with 4-λ WDM (one SOP per λ) without employing wavelength-selective devices; thus considerably relaxing the requirements on the lasers' frequency stability and removing the power consumption ascribed to thermal control of the passive optical filters. The second block covers such experimental validation, where two distinct transmission schemes are considered aiming 100 Gbps net-rate, that is 4×27 GBd (108 Gbps) and 4×32 GBd (128 Gbps) with hard-decision 7%- and 20%-overhead BCH FEC respectively. Three different system figures are discussed: receiver sensitivity, channel estimation robustness under fast polarization rotations, and power consumption plus complexity benchmarking.

The receiver sensitivity performance is studied in terms of uncoded

BER versus optical power at the receiver input for B2B configuration and C-band 2-km SSMF transmission. Below-FEC threshold demodulation is observed in all cases, proving successful communication employing fourth-order PolMux in IM/DD fiber-optic systems. Particularly for the 4×32-GBd case, ≈15-dB sensitivity penalty is measured with respect analogous WDM approach with ideal optical filtering. This number is theoretically argued, manifesting the strong base power penalty that that non-orthogonal PolMux entails.

By inducing polarization changes around one order of magnitude faster than expected in buried fiber and around four times faster than aerial-routed scenarios, $\approx 10^{-3}$ time-averaged BER is achieved with maximum instantaneous BER $\approx 8 \cdot 10^{-3}$ for the 4×32-GBd case at 25.5 dB of SNR. Despite successful below-FEC threshold demodulation, the instabilities in the instantaneous BER suggest using shorter period (<16000 symbols) between channel matrix updates.

Finally, a brief discussion on the power consumption and the DSP complexity is included. It is shown that in WDM scenarios where the carriers transport a single higher-speed serial data stream, the additional power consumption with respect to parallel optics is mainly caused by the extra driving current to overcome the sensitivity penalty. In the alternative scenario where every carrier belongs to a different user, the three extra ADCs (four in total) necessary to recover a single channel represent the highest fraction of the total power consumption. Along with the lasers driving, the power consumption in the latter case turns ≈37% higher than parallel optics. For its part, and due to its simplicity, the additional DSP as described in Section 4.2.5 constitutes 2.47% of the total consumption. Equivalently, up to fourth-order PolMux in IM/DD systems is possible by employing a Stokes-based set of algorithms implying as few as <1% extra summations and multiplications.

Future work

The present investigation conveys the following items:

- ISI cancellation : the accuracy of the polarization-tracking algorithm is hindered by ISI, thus enhancing current channel estimation to support multi-tap filters adaptation represents an intuitive solution [198]. The processing complexity and its impact on the power consumption must be considered.

- Polarization-tracking robustness and stability : the proposed algorithm requires reliable and consistent initialization when (i) the system operation is started and (ii) when the tracking is lost due to extremely-fast polarization changes. Data-aided channel estimation can be employed together with the presented algorithm for robustness increase [217]. Additionally, dynamic threshold adjustment in 4.14 (K) and shortening the time-period between channel matrix updates, can help with the precision of channel estimation and the BER stability.

Chapter 5

Conclusion

As the ever-evolving requirements in data throughput and connectivity keep pushing the existent electro-optical components plus transport networks to their functional limits, solutions facilitating efficient resource allocation gain proportional interest. In this line, DSP represents a powerful cost-effective tool, capable of concurrently improve spectral and power utilization through advance modulation and multiplexing schemes, coding, and electronic compensation of linear and non-linear impairments.

In this Thesis, we go beyond the state of the art by further exploiting the potential of DSP. We propose, develop, and experimentally validate for the first time in optical communications three advanced modulation and/or multiplexing techniques, each formulated according to the technical requirements of their specific scenarios. The target was to adapt the characteristics of the signal and the digital TX/RX features to better employ the available capacity. These three studies are initial steps towards the complete feasibility evaluation of the respective techniques for their use in actual systems.

In the following, we summarize the major findings and conclusions of each individual chapter.

5.1 (Ultra) long-haul networks

The need : currently, nonlineartities represent the major hindrance to increase the throughput-times-distance product in these scenarios. Customary treated as noise, it has been demonstrated that nonlinearities in >1000-km lumped-amplified uncompensated links show an additive Gaussian imprint on the constellation clusters; whereof we can leverage to apply

mature constellation-shaping techniques tracing capacity-achieving Gaussian distributions to boost the receiver sensitivity.

Proposal : SCM-PSM has been proposed as scalable and flexible coded modulation scheme with shaping capabilities featuring lower complexity in both transmitter and receiver sides than other alternatives. While this scheme has been long explored for wireless systems, it had not been demonstrated for optical communications.

Less than 8.4% information rate penalty for up to 12 bits per symbol is calculated while monotonically increasing the wellness of fit to normal distribution. The occasional ambiguous mapping for certain modulation orders (i.e. 6 bit/symbol) has been proved beneficial, resulting in larger average Euclidean distance among symbols than same-order QAM when the average power is equalized. This non-bijectivity in combination with the circular-symmetry of PSM also results in better defined power levels, whereby typical modulus-driven equalization techniques (e.g. RDE) base their operation on. This potentially improves the accuracy of the channel estimation. As far as decoding is concerned, we conclude that iterative demapping/decoding routines are necessary in order to achieve successful demodulation without strongly penalizing the effective rate; for which, nevertheless, lower coding gains than the analogous QAM are required.

Experimental proof of concept : successful decoding and demodulation is achieved after coherent detection of single-carrier DP 6-GBd 16-ary, 32-ary and 64-ary PSM constellations in B2B configuration and after EDFA-only uncompensated 240-km SSMF transmission. According to the SE results, 0.5 code-rate convolutional codes are shown highly inefficient despite the relatively similar EXIT profile with respect to PSM's; with more than 3 dB SNR gap to capacity. For their part, B2B measurements confirm the unsuitability of DPLLs for conducting carrier recovery at the low-SNR regions where shaping gain is achieved, as it caused the majority of the implementation penalties. Finally, the use of EM for cluster parametrization in order to feed the soft-based receiver with more accurate statistics proved $\approx$0.6 dB sensitivity improvement in highly non-linear transmission regime. The use of EM (or similar options) turns then interesting to allocate certain performance margins in those cases where the compliance with circularly symmetric Gaussian noise is not perfectly met. This is improvement is not related to the shaping gain.

Future : the results convey putting efforts in the code design and the robustness of carrier recovery; where multi-layer coding strategies and the use use of frequency pilot tones are concrete recommendations.

5.2 Metropolitan networks

The need : seeking for transparent data transport with high degree of connectivity, MAN networks progressively converge to all-optical reconfigurable mesh architectures. Besides the benefits, this class of network infrastructure poses challenges including pronounced over-filtering caused by manyfold OADM cascadability plus electronics, larger PMD, and extra OSNR degradation owed to the absence of intermediate regeneration. Consequently, reconfigurable software-driven sub-band/sub-carrier multiplexing is reasonably contemplated to increase the robustness against filtering effects and reduce the susceptibility to PMD without trading SE.

Proposal : C-MultiCAP has been conceived and proposed a scheme for sub-band de/multiplexing and de/modulation which groups together the multi-dimensionality of CAP and the spectral flexibility of MultiCAP for application in coherent communications.

We have seen that the use of $\geq$2 CAP dimensions per band gets relegated to, perhaps, implementing power-efficient multi-dimensional modulation formats of $\leq$5 order. This is due to the quadratic growth of total number of required filter taps in the corresponding de/multiplexing, and the strong increase on the PMEPR. This observation directly maps to the multi-band case, where the number of required filter taps is observed to increase approximately with the square of the product between the number of CAP dimensions and bands. In addition, the PMEPR may widely exceed DMT's for certain combinations of bands and CAP dimensions; all in all setting a pragmatic limit to both spectral granularity and dimensionality ($\approx$15-20% of the total bandwidth for two dimensions per band). The above observations base the decisions taken on the conception of the proposed C-MultiCAP system. We note that, by realizing early carrier recovery in the RX DSP (FPT technique is suggested) and forcing identical generation parameters in the four coherent dimensions, blind equalization taking advantage of conventional algorithms (e.g. CMA or RDE) is possible.

Experimental proof of concept : aiming for verifying the soundness of the assumptions supporting the proposed C-MultiCAP system, we im-

plement and transmit two different 2D-C-MultiCAP configurations over 225-km and 451-km EDFA-only SSMF links. The two generation schemes are: 3-band partitioning at 221.88 Gbps occupying 29.4-GHz optical bandwidth (≈7.5 b/s/Hz gross SE), and 5-band partitioning at 336 Gbps occupying 43-GHz optical bandwidth (≈7.8 b/s/Hz gross SE). We conclude that this combination of experimental conditions and generation parameters translates into notorious implementation penalties (≈5 dB and ≈8 dB for 221 Gbps and 336 Gbps respectively) and transmission penalties (≈5 dB after distance doubling for 221 Gbps) owed to the imperfect fulfillment of the basal assumptions. If the proposed equalization technique is to be used, narrower bands and looser inter-band spacing are solutions to improve the performance at the cost of more complexity and bandwidth. For its part, owed to CD-induced group-delay, FPT-based phase-noise estimation has been argued to be the main reason for transmission penalties and it shall be improved (e.g. spreading multiple tones across the spectrum) or alternative approaches.

On the other hand, the spectral and rate adaptability of C-MultiCAP have been demonstrated by achieving >300 Gbps in a system with less than 6-GHz end-to-end bandwidth; as well as the ease of reconfigurability, allowing for >100-Gbps line rate increase only adjusting the TX/RX DSP. Finally, recovery and equalization of multiple correlated CAP dimensions has been shown, proving the plausibility of their joint exploitation for implementing multi-dimensional modulation formats.

Future : generation parameters, phase noise estimation, and equalization need to be perfected for ensuring robust demodulation and higher sensitivities (approximately from 3 to 5 dB better is possible). The reduction of processing complexity should be pursued.

5.3 Short-haul networks

The need : with short-haul subsystems mainly driven by cost and power consumption, WDM and parallel optics have been preferred methods to keep pace with throughput growth due to the apparent de/multiplexing simplicity. Nevertheless, as this throughput demand grows steadily, the scaling efficiency gets jeopardized. For instance, an excessive number of wavelengths requires strict wavelength stability control and makes packaging challenging; while too many fiber-pairs would likewise suffer from packaging and fiber-transceiver interfacing issues. This outlook encourages

the incorporation of polarization-multiplexing to make better use of the available bandwidth and/or space.

Proposal : we have introduced a polarization-multiplexed technique enabling simultaneous transmission of $\leq$4 distinct SOPs in IM/DD fiber-optic communication systems, and usable alongside or in place of existing technologies (e.g. WDM or multilevel modulation).

During the initial explanation of the fundamentals, the limitation to 4 as the maximum number of demultiplexable SOPs is argued; as well as their optimum orientation with respect to the receiver's polarizers to improve the robustness and simplicity of the channel estimation and minimizing noise enhancement. This is followed by the elaboration on one possible RX DSP chain for blind low-complexity quasi-continuous SOP tracking and signal demultiplexing. Finally, exploiting the possibility to go beyond traditional dual-polarization for SE doubling, we propose $\geq$3-SOP configurations as alternative demultiplexing method in other multiplexing technologies (e.g. WDM); thereby enabling the interchange of the respective apparatus.

Experimental proof of concept : based on the last-mentioned idea, we conduct its experimental validation in a 4-λ WDM system, where wavelength demultiplexing is performed without the need for optical filters. This potentially relaxes the requirements on the lasers' stability and avoids the power consumption associated to the thermal control of such filters. Targeting 100-Gbps net rate, two WDM systems are evaluated: 4×27 GBd (108 Gbps gross) and 4×32 GBd (128 Gbps gross) considering 7%- and 20%-overhead FEC respectively.

Besides achieving successful wavelength demultiplexing and signal demodulation, the results in terms of receiver sensitivity evince the strong base power penalty that that non-orthogonal polarization-multiplexing ($\geq$3-SOP) entails ($\approx$15-dB sensitivity penalty with respect to ideally filtered WDM). However, below-FEC average BER performance is observed even under abnormally fast polarization changes; still motivating the application of the proposed DSP for other configurations (e.g. standard dual-polarization OOK). In terms of power consumption, the proposed system consumes $\approx$37% more than parallel optics, from which only 2.47% belongs to the software; or equivalently, as few as $\approx$1% extra summations and multiplications.

Future : main recommendations relate to improving the robustness of the

polarization tracking algorithm; certainly susceptible to ISI and imprecise initialization.

Appendix A

SCM-PSM 1D gaussianity

This appendix provides comprehensive and complementary study on the convergence evolution of SCM-PSM output constellations' distribution towards Gaussianity. The metrics used permit adaptation to any reference distribution (not necessarily Gaussian), hence constituting a straight and generalizable method for analyzing the shaping gain evolution for any given memoryless channel and comparing the results among various modulation alternatives.

Tools

The evaluation of PSM constellations' distribution is performed through numerical and visual analysis of the cdf of the individual quadrature components in the absence of noise. The visual test is performed with a statistical tool named quantile-quantile (Q-Q) plot, which is a graphical method for the diagnosis of fitting deviations between two probability distributions based on percentile comparison. These percentiles are measured or estimated from the cdf of the respective random variables and then plotted against each other in an unequivocal scatter plot. Given that we want to check upon the Gaussianity of certain data sample, one of the probability distributions acts as the reference model with a theoretical normal profile $N(\mu, \sigma)$, while the other corresponds to the cdf of PSM constellations' quadrature projections. This constitutes the most basic plotting format and it is referred as one-sample Q-Q plot. It is noteworthy that, since the reference distribution is continuous and known, all possible quantiles are calculable through simple cdf inversion. This means that any value from the discontinuous cdf can be directly represented thence avoiding inaccurate

data interpolation. The descriptive parametric equations of the one-sample Q-Q plot are given by:

$$\begin{aligned} Y_i &\equiv \text{Data quantiles} = F^{-1}_{DATA}(p_i) \\ X_i &\equiv \text{Reference quantiles} = F^{-1}_{REF}(p_i) \end{aligned} \tag{A.1}$$

where F^{-1}_{DATA} and F^{-1}_{REF} represent the quantile functions of the cdfs for the data and the reference model respectively, and $p_i \in (0, 1)$ determines the order of the ith quantile. In order to normalize the axes' scales, the mean (μ) and variance (σ^2) of the theoretical normal reference are equalized to the estimated mean and variance of PSM quadrature components for every N. In that case, the distributions being analyzed will be identical if $\forall i$:

$$F^{-1}_{DATA}(p_i) = F^{-1}_{REF}(p_i) \tag{A.2}$$

The evaluation of the resulting curves will provide qualitative information on the distribution shapes, tail behavior, mean/variance deviations, outliers and symmetry unbalances. However, in order to obtain numeric evaluation of the goodness of fit between the empirical and theoretical cdf, certain statistical criteria must be associated with each case of study. In this research, the fitting variance (σ_{FIT}) and the Kolmogorov-Smirnov statistic factor (D) are used as quantitative indicators. The mathematical expressions:

$$\begin{aligned} \sigma^2_{FIT} &= \mathrm{E}\left[|F^{-1}_{DATA}(p_i) - F^{-1}_{REF}(p_i)|^2\right] \\ D &= \sup\left[|F^{-1}_{DATA}(p_i) - F^{-1}_{REF}(p_i)|\right] \end{aligned} \tag{A.3}$$

where $\sigma^2_{FIT}, D \in [0, \infty)$. Whereas σ^2_{FIT} gives information about the average Gaussian purity of the empirical distribution, D shows possible outliers and symmetry disturbances. Both figures of merit denote good agreement for values approaching zero.

Results and discussion

Figure A.1 shows the visual test based on Q-Q plotting for $N \in \{4, 6, 9\}$. The corresponding cross-sectional probability mass functions of the constellations are included for the reference. Figure A.2 displays the numerical results of goodness of fit based on the normal fitting variance and the one-sample Kolmogorov–Smirnov (KS) statistic for both QAM and PSM up to

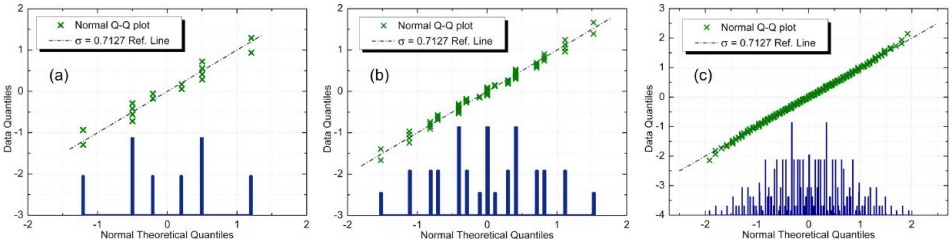


Figure A.1: Q-Q plot Gaussianity test on the quadrature projections of SCM-PSM constellations for $N = 4$ (a), $N = 6$ (b) and $N = 9$ (c).

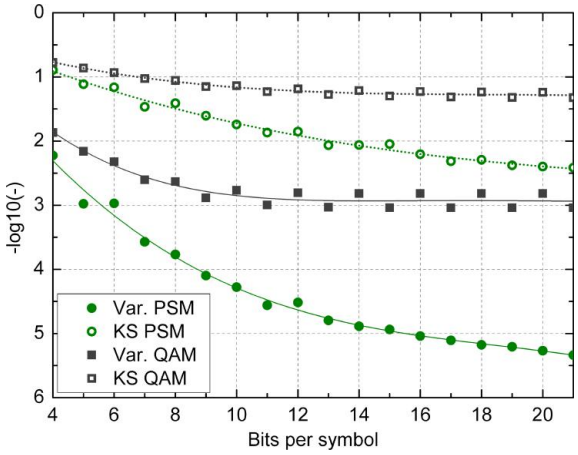


Figure A.2: Gaussian-fitting variance and Kolmogorov-Smirnov factor of QAM and PSM for $\{N \in \mathbb{N}, 4 \leq N \leq 21\}$.

$N = 21$. The fitting variance and KS statistic are chosen because of their complementarity, simplicity of calculation, and ease of understanding. Tendency lines are also presented for reference purposes.

It is observed in Figure A.1 that the dispersion of the Q-Q points around the reference normal distribution decreases as the modulation order increases. As N increases good fitting is observed around quantile zero, where the average deviation from Gaussianity is dominated by disagreements on the tails (see Figure A.1c). Notice that only one cross-section of the power-normalized constellation is analyzed, hence the probability distribution of the Gaussian reference that yields one-to-one quantile scaling is given by $\mu = 0$ and $\sigma^2 = 0.5$ ($\sigma \approx 0.7$).

Figure A.2 shows that the average Gaussianity for every N and modulation format is determined by the KS statistic with some exceptions in

SCM-PSM comprising N6,8,14,15,18. The difference in the fitting variance between PSM and QAM exceeds one order of magnitude for $N = 9$ and two orders for $N = 16$. For low number of bits per symbol, the quantization error governs the deviation with respect to Gaussianity. This fact explains the convergence to normal distribution of QAM for $N \in [0, 10]$, and the fast convergence of PSM (two orders of magnitude within $N = 4$ and $N = 11$). Once the quantization error stops dominating, QAM shows flat tendency and PSM shows $\approx$0.65 dB/N convergence rate to unit normal distribution. Note that owed to the square-distribution of QAM constellations' for N even, σ^2_{FIT} is 2 dB higher than cross-distributed orders, i.e. N odd. The monotonic decrease of the fitting curve for PSM which partially proves the asymptotic tendency to Gaussianity.

Central limit theorem

Because in SCM signal generation is carried out via the superposition of N data-independent sequences, the convergence to Gaussianity can be readily tested through the evaluation of the CLT. Particularly in PSM, the phase-shift induced by the complex weights before superposition makes the random independent data streams not be identically distributed, hence preventing the strict application of the traditional CLT definition. In this case, Lyapunov condition can be evaluated instead.

The Lyapunov condition states that, given a sequence of independent but not necessarily identically distributed random variables X_i, with finite mean μ_i and variance σ_i^2, and

$$r_n^{2+\epsilon} = \sum_{i=1}^{n} \mathrm{E}\left[|X_i - \mu_i|^{2+\epsilon}\right] \tag{A.4}$$

if for some $\epsilon > 0$

$$\lim_{n\to\infty} \frac{r_n^{2+\epsilon}}{s_n^{2+\epsilon}} = 0 \tag{A.5}$$

where

$$s_n^2 = \sum_{i=1}^{n} \sigma_i^2 \tag{A.6}$$

then the CLT holds.

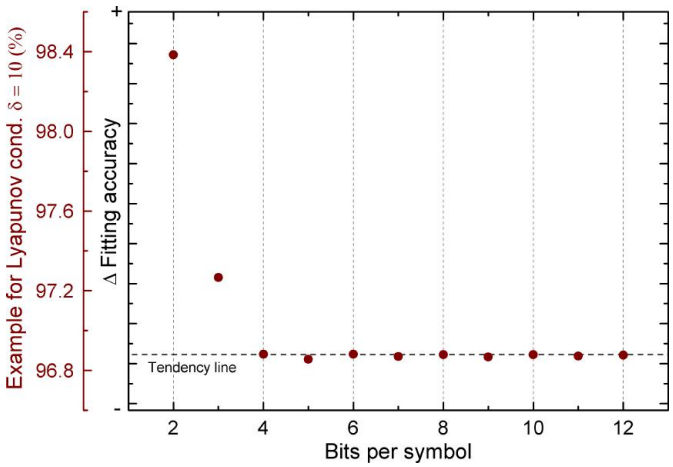


Figure A.3: Conceptual illustration of the proposed quadruple polarization system for IM/DD optical data links.

Figure A.3 shows the evaluation of A.5 versus the number of added random variables n (i.e. number of bits/sym N) for $\epsilon = 10$ (arbitrary). For the ease of visualization, the increment magnitude is show instead; meaning that, if the curve stabilizes, the convergence to zero (Gaussianity) is relentless and invariant. The results prove positive convergence up to $N = 12$ with constant increment tendency, presumably signifying that CLT applies and thus the asymptotic convergence to Gaussianity of the marginal quadrature distributions.

Appendix B

Memoryless channel-capacity

In the optical communications society, the word *capacity* is often employed alluding to two related but distinct concepts. The first one is the traffic volume or amount of information per unit time, which is referred to as *throughput* across the entire document. And secondly, as a characteristic parameter of any noisy transmission medium or channel, by which the maximum amount of information that the observed output can ever give about the input gets quantified. We use *capacity* for the latter meaning. In the following, we provide the mathematical fundamentals behind *capacity* calculation for memoryless channels.

In information-theory terms, the capacity of a given a noisy channel is expressed as the supremum of the mutual information across all possible input distributions:

$$C \triangleq \sup_{p_X} I(X;Y) \tag{B.1}$$

where p_X is the input signal's pdf and the mutual information is given by:

$$I(X;Y) = I(Y;X) \triangleq H(Y) - H(Y|X) = H(X) - H(X|Y) \tag{B.2}$$

where $H(\bullet)$ and $H(\bullet|\bullet)$ are the entropy and the conditional entropy operators respectively[1]. Figure B.1 intuitively illustrates the relation among

[1]For simplicity, no distinction is made between discrete and differential entropy concerning the operator's symbol.

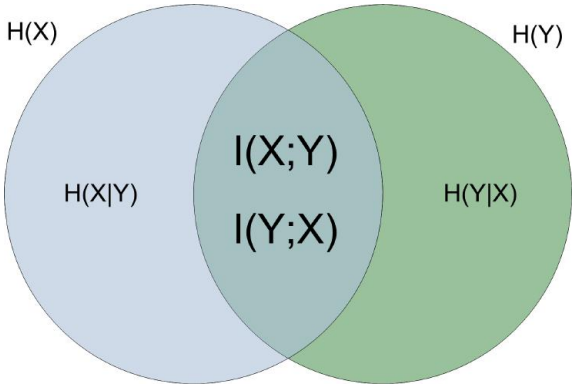


Figure B.1: Venn diagram relating various information measuremets.

important measurements of information, making $I(X;Y)$=$I(X;Y)$ apparent. Considering continuous complex input/output signals:

$$H(X) \triangleq -\int_{\mathbb{C}} p_X(x) \log_2 p_X(x) dx \tag{B.3}$$

$$H(X|Y) \triangleq -\iint_{\mathbb{C}} p_{X,Y}(x,y) \log_2 p_{X|Y}(x|y) dxdy \tag{B.4}$$

Substituting B.3 and B.4 into B.2 and leveraging Bayes' theorem[2]:

$$I(X;Y) \triangleq \iint_{\mathbb{C}} p_{X,Y}(x,y) \log_2 \frac{p_{X,Y}(x,y)}{p_X(x) p_Y(y)} dxdy \tag{B.5}$$

or alternatively, expressing B.5 as a function of the input distribution and the channel's transition probability:

$$I(X;Y) \triangleq \iint_{\mathbb{C}} p_X(x) p_{Y|X}(y|x) \log_2 \frac{p_{Y|X}(y|x)}{\int_{\mathbb{C}} p_X(x) p_{Y|X}(y|x) dx} dxdy \tag{B.6}$$

Discrete-input continuous-output

If the input to the channel is discrete with finite alphabet Υ, B.6 becomes:

$$I(X;Y) \triangleq \sum_{x \in \Upsilon} p_X(x) \int_{\mathbb{C}} p_{Y|X}(y,x) \log_2 \frac{p_{Y|X}(y|x)}{\sum_{x \in \Upsilon} p_X(x) p_{Y|X}(y|x)} dy \tag{B.7}$$

[2] $p_{X,Y}(x,y) = p_{X|Y}(x|y) p_Y(y)$

The resulting mutual information is called *constrained capacity*, making direct allusion to the fact that the optimization is performed on a subset of possible input distributions. Note that given a continuous-output channel, conditioning the mutual information optimization to finite-alphabet input unavoidably results in a capacity lowerbound.

The link between the purely statistical expression B.7 and physically meaningful parameters turns apparent when the finite input alphabet is related to a particular M-ary constellation, e.g. the 16 points of 16-QAM or the 4 points of QPSK. Upon such interpretation, the relation between *constrained capacity* and the symbols' relative position (present in $p_{Y|X}(y|x)$) and/or their probability of occurrence ($p_X(x)$) gets intuitively revealed, as well as the basis of *constellation shaping*. Accordingly, the following inequality holds in scenarios where the input power is limited:

$$\sum_{x \in \Upsilon} |x|^2 p_X(x) \leq P \tag{B.8}$$

In the case where the channel's noise is zero-mean complex AWG, $p_{Y|X}(y|x)$ is given by:

$$p_{Y|X}(y|x) = p_Z(y - x) \tag{B.9}$$

where

$$p_Z(z) = \frac{1}{\pi\sigma^2} e^{\frac{|z|^2}{\sigma^2}} \tag{B.10}$$

where σ^2 is the total complex-noise power. Because under AWGN the integral in B.7 is of the kind $\int_{-\infty}^{\infty} e^{-x^2} f(x)dx$, the Gauss-Hermite quadrature numerical approximation shall be employed to accelerate the calculation:

$$\int_{-\infty}^{\infty} e^{-x^2} f(x)dx \approx \sum_{i=1}^{n} \omega_i f(x_i) \tag{B.11}$$

where n is the number of sample points used and ω_i is given by:

$$\omega_i = \frac{2^{n-1} n! \sqrt{\pi}}{n^2 \left[Q_{n-1}(x_i)\right]^2} \tag{B.12}$$

where $Q_n(x)$ is the Hermite polynomial. In this investigation we employed a 16-point (n) approximation as the optimum quadrature for polynomials of degree 31 or less, widely sufficing for the accurate computation of capacity curves.

AWGN channel capacity : Under AWGN, B.2 reduces to:

$$\begin{aligned} I(X;Y) &= H(Y) - H(X+Z|X) \\ &= H(Y) - H(Z) \end{aligned} \tag{B.13}$$

inserting B.12 into B.13:

$$I(X;Y) = H(Y) - \log_2 \pi e \sigma^2 \tag{B.14}$$

Because the differential entropy maximizes when the random variable describes a Gaussian distribution, and knowing that the noise and the input are statistically independent:

$$I(X;Y) \leq \log_2 \pi e(P + \sigma^2) - \log_2 \pi e \sigma^2 \tag{B.15}$$

where P is the input signal power. Then resolving for the maximum, the 2D (complex) continuous-input continuous-output channel capacity is obtained:

$$C = \log_2 \left(1 + \frac{P}{\sigma^2}\right) \tag{B.16}$$

Appendix C

Frequency pilot tone

The principle of FPT-based frequency offset and phase noise estimation is very intuitive. Let the noiseless transmitted field of the FPT in one polarization be expressed as:

$$E_t(t) = E_{t;0} e^{i2\pi(f_{FPT}+f_C)t+\phi_C(t)} \tag{C.1}$$

where f_C and $\phi_C(t)$ are the carrier frequency and phase respectively, f_{FPT} is a user-defined intermediate frequency (DC is also considered), and $E_{t;0}$ is the constant-valued field's amplitude function, in turn dependent on the FPT insertion technique. After ideal transmission and coherent detection, the field is given by:

$$\begin{aligned} E_d(t) =& E_{d;0} e^{i2\pi(f_{FPT}+f_C-f_{LO})t+\phi_C(t)-\phi_{LO}(t)} \\ =& E_{d;0} e^{i2\pi(f_{FPT}+\Delta f)t+\Delta\phi(t)} \end{aligned} \tag{C.2}$$

where f_{LO} and $\phi_{LO}(t)$ are the frequency and phase of the LO, and $E_{d;0}$ is the constant received field's amplitude. Because f_{FPT} is known beforehand, a simple evaluation of the argument of C.2 yields Δf, the frequency offset after coherent beating; and $\Delta\phi(t)$, the combined phase-noise function of both transmitter and receiver lasers. The frequency tone is thus able to blindly and continuously track the evolution of the phase fluctuations experienced by the electric field at the exact frequency where it is inserted. In the receiver, such phase distortions are uniformly corrected for across the entire detected spectrum; which includes the co-propagating data signal.

However, the accuracy of this technique relies on two important assumptions made on the above mathematical elaboration and whose validity weakens in real transmission scenarios:

- Absence of noise plus the perfect isolation of the FPT upon reception; which not being the case, requires careful optimization of the PSR, and allocating enough guard-band around the FPT. Both directly related to the phase estimation fidelity.

  By increasing the PSR, the influence of the noise in the phase estimation becomes weaker, while excessively high PSR results in OSNR degradation for the information signal and quantization noise when digital FPT generation is employed. The optimum ratio depends on the system's noise power, the signal bandwidth, the phase noise characteristics, and the properties of the filter employed for FPT isolation. Experimental characterization and empirical observations in various demonstrations [90, 157, 158] indicate optimum PSR of ≈-15 dB.

  Then FPT guard-banding, dictating the trade-off between total effective rate and the maximum frequency excursions that can be tracked within affordable performance degradation. Given a fixed observation time, time-varying phase fluctuations show as signal broadening in the frequency domain; therefore, proportional margin between the FPT and the closest frequency components of the information signal must be allocated so as not to incur spectral overlapping, and improve FPT-signal isolation ratio. The particular guard-band arrangement depends on the characteristics of the spectral tails of the neighboring signal, and on possible constraints on the filter design.

- The phase distortion that the FPT reads is common to all spectral components; which is false in the presence of any mechanism inducing certain degree of group delay, such as some analog filters or CD. For instance, CD-induced walk-off effect will cause a frequency-dependent decorrelation of the phase vectors by which neighboring components to the FPT will exhibit better performance than those far apart [156]. Possible solutions to alleviate this imprecision include the use of multiple FPTs in different frequency regions [122], enabling more dedicated phase-noise compensation (for instance, one FPT per sub-band in C-MultiCAP systems) and/or measurement noise averaging; and placing the FPTs such that the maximum distance to all transmitted spectral components and other FPTs is minimized.

With regard to implementation, FPT insertion can be realized in three alternative ways. Figure C.1 shows a conceptual illustration of the corresponding output spectra:

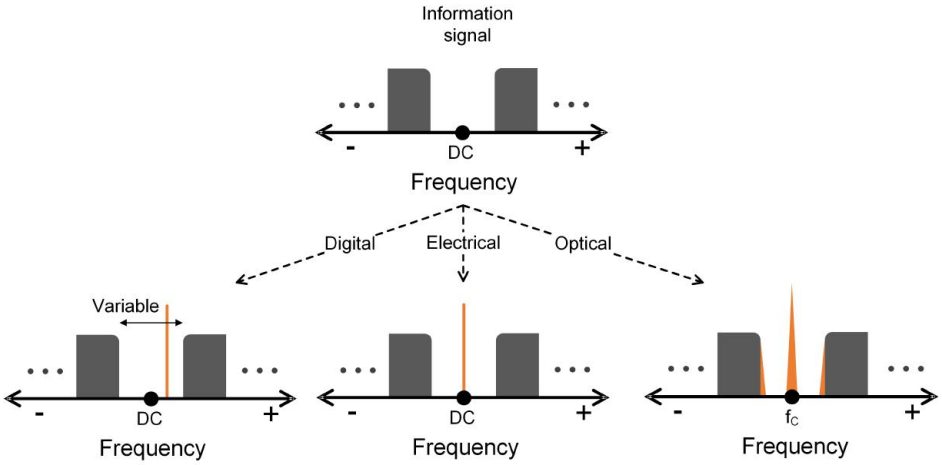


Figure C.1: On the top, FPT-less information signal. On the bottom from left to right, conceptual representation of the ouput spectra of the three main techniques for FPT insertion: digital, electrical and optical.

- *Digital (e.g. [46]) :* the FPT is inserted at the end of the transmitter's DSP. Depending on the FPT allocation, the insertion stage may vary between prior or posterior to high-frequencies enhancement pre-distortion to minimize the manipulation of the tone before digital-to-analog conversion. This approach permits highly accurate center-frequency allocation, readily allowing for radio-frequency FPTs and hence a desirable null DC component at the expense of quantization noise. Care should be taken to ensure an integer number of period within DAC's repetition memory so as not to incur in windowing effects that degrade the later phase estimation.

- *Electrical analog (e.g. [90]):* certain DC level is intentionally induced during electro-optical conversion. Precise PSR control, distortion-less FPT insertion, and hardware minimization is obtained at the expense of higher average power plus consumption, and potential non-linear signal distortions according to the modulator's response.

- *Optical analog (e.g. [155]) :* in this method a copy of the carrier is inserted without intermediate DSP or electro-optical interfaces. Whereas the FPT phase is not distorted, additional pieces of optical hardware are necessary.

Appendix D

DSP schema for 4-SOP IM/DD transmission system

```
% INTENSITY SIGNAL MANIPULATION

% Signal conditioning

Resampling % Outputs 2 samples per symbol
Filtering % Low-pass filtering for SNR improvement


% Front-end correction

Skew % time skew compensation
Responsivity % Photodioide's responsivity equalization
DC Compensation % DC offset correction



% STOKES PARAMETER MANIPULATION

Intensity to Stokes conversion % Stokes parameter generation
% based on 1.10
Timing Recovery % time-domain Gardner error estimation
% and decimation


% SOP tracking

Zero-power discriminator % Trivial zero-power case is detected,
```

```
% demodulated and removed.
Stokes parameters normalization % S1 to S3 are normalized by S0
Initialization of coordinate system and decision thresholds % In
% the absence of initialization algorithm (e.g. training sequence),
% the coordinate system may be initialized with the identity matrix.

while not(convergence)
Estimate Stokes parameter (4.15)

    switch estimated Stokes parameter
        case +/−1
            Calculate the sign−dependent error signal (4.14)
            Update the coordinate system (4.16).
            Update decision thresholds % dependent on noise,
            % modulation format, and transmitted plus detected SOPs
            Stokes parameter index = Stokes parameter index;
            symbol to analyze = symbol to analyze + 1;
        case 0
            Stokes parameter index = Stokes parameter index + 1;
            symbol to analyze = symbol to analyze;
    end

end


% Channel distortion compensation
Transpose estimated channel matrix (4.17)
Compensate distortion (4.2)


% Demultiplexing
SOP demultiplexing (e.g. 4.18)


% Demodulation
Demodulation and error counting per channel
```

List of Acronyms

1D one-dimensional

2D two-dimensional

3D three-dimensional

ADC analogue-to-digital converter

AMF advanced modulation formats

ASE amplified spontaneous emission

ASIC application-specific integrated circuit

aSOP aggregate SOP

AWGen arbitrary waveform generator

AWGN additive white Gaussian noise

B2B back-to-back

BCJR Bahl-Cocke-Jelinek-Raviv

BER bit error rate

BICM bit-interleaved coded modulation

BPSK binary phase-shift keying

CAP carrierless amplitude/phase modulation

CCDF complementary cumulative distribution function

CD chromatic dispersion

cdf cumulative distribution function

CER constellation expansion ratio

CLT central limit theorem

CMA constant-modulus algorithm

C-MultiCAP coherent multi-band carrierless amplitude/phase modulation

DAC digital-to-analogue converter

DC direct current

DD direct-detection

DFB distributed feedback

DGD differential group delay

DMT discrete multitone

DP dual-polarization

DPLL digital phase-locked loop

DSO digital storage oscilloscope

DSP digital signal processing

ECL external-cavity laser

EDFA Erbium-doped fiber amplifier

EM expectation maximization

ETDM electrical time-division multiplexing

EXIT extrinsic information transfer

FEC forward error correction

FIR finite impulse response

FPT frequency pilot tone

IBI inter-band interference

ICI inter-carrier interference

ID iterative decoding-demapping

IDF inverse dispersion fiber

IM intensity-modulation

IPM iterative polar modulation

ISI inter-symbol interference

LAN local-area network

LC left-circular

LDPC low-density parity-check

LH long-haul

LO local oscillator

MAC medium access control

MAN Metropolitan-area network

MAP maximum a posteriori probability

MI mutual information

MIMO multi-input multi-output

MLSE maximum-likelihood sequence estimation

MMA multi-modulus algorithm

MMSE minimum mean square error

MoG mixture of Gaussian

MultiCAP multi-band carrierless amplitude/phase modulation

MZM Mach-Zehnder modulator

NLMS normalized least mean square

NMSE normalized mean square error

NRZ non-return-to-zero

NZDSF non-zero dispersion-shifted fiber

OADM optical add-drop multiplexer

OFDM orthogonal frequency-division multiplexing

OOK on-off keying

OSNR optical signal-to-noise ratio

OTDM optical time-division multiplexing

P2P point-to-point

PAM pulse-amplitude modulation

PAPR peak-to-average power ratio

PBS polarization beam splitter

pdf probability density function

PDL polarization-dependent loss

PMD polarization-mode dispersion

PMEPR peak-to-mean envelop power ratio

pmf probability mass function

PM-QPSK polarization-multiplexed quadrature phase-shift keying

PolC polarization controller

PolMux polarization multiplexing

PolSK polarization-shift keying

PPG pulse pattern generator

PRBS pseudo-random bit sequence

PSCF pure silica core fiber

PSD power spectral density

PSM phase-shifted mapping

PSR pilot-to-signal ratio

QAM quadrature amplitude modulation

QMSE quantization mean-square error

QPSK quadrature phase-shift keying

Q-Q quantile-quantile

RC right-circular

RDE radius-directed equalizer

RRC root-raised-cosine

RX receiver

RZ return to zero

SCM superposition coded modulation

SE spectral efficiency

SER symbol error rate

SISO soft-input soft-output

SLAF super-large-area fiber

SNR signal-to-noise ratio

SNRb signal-to-noise ratio per bit

SNRs signal-to-noise ratio per symbol

SOP state of polarization

SPM self-phase modulation

SSMF standard single mode fiber

TX transmitter

ULAF ultra-large-area fiber

ULH ultra long-haul

VOA variable optical attenuator

WDM wavelength-division multiplexing

Bibliography

[1] K. Kao and G. Hockham, "Dielectric-fibre surface waveguides for optical frequencies," *Proceedings of the Institution of Electrical Engineers*, vol. 113, no. 7, p. 1151, 1966.

[2] H. Schoeneich and P. A. Hoeher, "Adaptive interleave-division multiple access-a potential air interface for 4G bearer services and wireless LANs," in *2004 International Conference on Wireless and Optical Communications and Networks (WOCN)*, Jun. 2004.

[3] T. Wo and P. Hoeher, "Superposition mapping with application in bit-interleaved coded modulation," 2010, pp. 1–6.

[4] D. Hao and P. A. Hoeher, "Superposition modulation with reliability-based hybrid detection," in *2010 6th International Symposium on Turbo Codes & Iterative Information Processing*. IEEE, Sep. 2010, pp. 280–284.

[5] T. Wo and P. A. Hoeher, "A universal coding approach for superposition mapping," in *2010 6th International Symposium on Turbo Codes & Iterative Information Processing*. IEEE, Sep. 2010, pp. 314–318.

[6] T. Wo, M. Noemm, D. Hao, and P. A. Hoeher, "Iterative Processing for Superposition Mapping," *Journal of Electrical and Computer Engineering*, vol. 2010, pp. 1–13, 2010.

[7] P. Hoeher and T. Wo, "Superposition modulation: myths and facts," *IEEE Communications Magazine*, vol. 49, no. 12, pp. 110–116, Dec. 2011.

[8] M. Noemm, A. Mourad, and P. A. Hoeher, "Superposition modulation with irregular convolutional coding," in *2012 IEEE Global Communications Conference (GLOBECOM)*. IEEE, Dec. 2012, pp. 2346–2350.

[9] D. Hao and P. A. Hoeher, "A low-complexity tree search detection algorithm for superposition modulation," in *2012 7th International Symposium on Turbo Codes and Iterative Information Processing (ISTC)*. IEEE, Aug. 2012, pp. 145–149.

[10] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-Model of Fiber Non-Linear Propagation and its Applications," *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 694–721, Feb. 2014.

[11] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Pulse collision picture of inter-channel nonlinear interference in fiber-optic communications," *Journal of Lightwave Technology*, vol. PP, no. 99, pp. 1–1, 2015.

[12] "Cisco visual networking index: Forecast and methodology, 2014–2019," Cisco, Tech. Rep., 2015.

[13] "Technology Options for 400G Implementation," Jul. 2015. [Online]. Available: http://www.oiforum.com/wp-content/uploads/OIF-Tech-Options-400G-01.0.pdf

[14] C. Colombo and S. Morganti, "Reconfigurable DP-16QAM/QPSK transponders for the transmission of coherent 200G carriers in a flexgrid super-channel arrangement," in *2014 19th European Conference on Networks and Optical Communications - (NOC)*. IEEE, Jun. 2014, pp. 179–184.

[15] A. Lord, P. Wright, and A. Mitra, "Core Networks in the Flexgrid Era," *Journal of Lightwave Technology*, vol. 33, no. 5, pp. 1126–1135, Mar. 2015.

[16] J. D. Reis, M. Garrich, D. M. Pataca, J. C. M. Diniz, V. N. Rozental, L. H. H. Carvalho, E. C. Magalhaes, U. Moura, N. G. Gonzalez, J. R. F. Oliveira, and J. C. R. F. Oliveira, "Flexible optical transmission systems for future networking," in *2014 16th International Telecommunications Network Strategy and Planning Symposium (Networks)*. IEEE, Sep. 2014, pp. 1–6.

[17] J. S. Eng, "What's Next for 100G and 400G Client Optics?" in *2014 Optical Fiber Communication Conference and the National Fiber Optic Engineers Conference. Market Watch Panel 4.*, 2014.

[Online]. Available: http://www.ofcconference.org/library/images/ofc/2014/Market%20Watch%20and%20SPS/2-JulieEng.pdf

[18] "Spectral grids for WDM applications: DWDM frequency grid," Feb. 2012. [Online]. Available: https://www.itu.int/rec/T-REC-G.694.1-201202-I/en

[19] "Implementation Agreement for Generation 2.0 100G Long-Haul DWDM transmission Module - Electromechanical (Gen.2 MSA-100GLH)," Aug. 2013. [Online]. Available: http://www.oiforum.com/public/documents/OIF-MSA-100GLH-EM-02.0.pdf

[20] "Forward error correction for submarine systems," Oct. 2000. [Online]. Available: https://www.itu.int/rec/T-REC-G.975-200010-I/en

[21] P. P. Mitra and J. B. Stark, "Nonlinear limits to the information capacity of optical fibre communications." *Nature*, vol. 411, no. 6841, pp. 1027–30, Jun. 2001.

[22] "Submarine telecoms industry report," Terabit Consulting, Tech. Rep., 2014.

[23] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity Limits of Optical Fiber Networks," *Journal of Lightwave Technology*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[24] A. D. Ellis, J. Zhao, and D. Cotter, "Approaching the Non-Linear Shannon Limit," *J. Lightwave Technol.*, vol. 28, no. 4, pp. 423–433, 2010.

[25] R. I. Killey and C. Behrens, "Shannon's theory in nonlinear systems," *Journal of Modern Optics*, vol. 58, no. 1, pp. 1–10, Jan. 2011.

[26] G. Bosco, P. Poggiolini, A. Carena, V. Curri, and F. Forghieri, "Analytical results on channel capacity in uncompensated optical links with coherent detection." *Optics express*, vol. 19, no. 26, pp. B438–49, Dec. 2011.

[27] ——, "Analytical results on channel capacity in uncompensated optical links with coherent detection: erratum," *Optics Express*, vol. 20, no. 17, p. 19610, Aug. 2012.

[28] A. Mecozzi and R.-J. Essiambre, "Nonlinear Shannon Limit in Pseudolinear Coherent Systems," *Journal of Lightwave Technology*, vol. 30, no. 12, pp. 2011–2024, Jun. 2012.

[29] M. Secondini, E. Forestieri, and G. Prati, "Achievable Information Rate in Nonlinear WDM Fiber-Optic Systems With Arbitrary Modulation Formats and Dispersion Maps," *Journal of Lightwave Technology*, vol. 31, no. 23, pp. 3839–3852, Dec. 2013.

[30] R. Dar, M. Shtaif, and M. Feder, "New bounds on the capacity of the nonlinear fiber-optic channel." *Optics letters*, vol. 39, no. 2, pp. 398–401, Jan. 2014.

[31] A. Morea, J. Renaudier, T. Zami, A. Ghazisaeidi, and O. Bertran-Pardo, "Throughput Comparison Between 50-GHz and 375-GHz Grid Transparent Networks [Invited]," *Journal of Optical Communications and Networking*, vol. 7, no. 2, p. A293, Dec. 2014.

[32] N. Amaya, M. Irfan, G. Zervas, K. Banias, M. Garrich, I. Henning, D. Simeonidou, Y. R. Zhou, A. Lord, K. Smith, V. J. F. Rancano, S. Liu, P. Petropoulos, and D. J. Richardson, "Gridless optical networking field trial: flexible spectrum switching, defragmentation and transport of 10G/40G/100G/555G over 620-km field fiber." *Optics express*, vol. 19, no. 26, pp. B277–82, Dec. 2011.

[33] S. Frisken, G. Baxter, D. Abakoumov, Hao Zhou, I. Clarke, and S. Poole, "Flexible and grid-less wavelength selective switch using LCOS technology," pp. 1–3, 2011.

[34] P. J. Winzer, "High-Spectral-Efficiency Optical Modulation Formats," *Journal of Lightwave Technology*, vol. 30, no. 24, pp. 3824–3835, Dec. 2012.

[35] F. Hauske, *The Importance of Digital Signal Processing in High Speed Optical Receivers: Equalization, Impairment Compensation and Performance Monitoring*, ser. Forschungsberichte aus der Kommunikationstechnik. Köster, 2013.

[36] E. Temprana, E. Myslivets, B. P.-P. Kuo, L. Liu, V. Ataie, N. Alic, and S. Radic, "Overcoming Kerr-induced capacity limit in optical fiber transmission," *Science*, vol. 348, no. 6242, pp. 1445–1448, Jun. 2015.

[37] X. Li, X. Chen, G. Goldfarb, E. Mateo, I. Kim, F. Yaman, and G. Li, "Electronic post-compensation of WDM transmission impairments using coherent detection and digital signal processing," *Optics Express*, vol. 16, no. 2, p. 880, Jan. 2008.

[38] D. J. Rezende, S. Mohamed, and D. Wierstra, "Stochastic Backpropagation and Approximate Inference in Deep Generative Models," Jan. 2014.

[39] L. Zhu and G. Li, "Nonlinearity compensation using dispersion-folded digital backward propagation." *Optics express*, vol. 20, no. 13, pp. 14 362–70, Jun. 2012.

[40] D. Rafique, M. Mussolin, M. Forzati, J. Må rtensson, M. N. Chugtai, and A. D. Ellis, "Compensation of intra-channel nonlinear fibre impairments using simplified digital back-propagation algorithm." *Optics express*, vol. 19, no. 10, pp. 9453–60, May 2011.

[41] T. Hoshida, L. Dou, W. Yan, L. Li, Z. Tao, S. Oda, H. Nakashima, C. Ohshima, T. Oyama, and J. C. Rasmussen, "Advanced and feasible signal processing algorithm for nonlinear mitigation," in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013*. Optical Society of America, Mar. 2013, p. OTh3C.3.

[42] L. B. Du and A. J. Lowery, "Improved single channel backpropagation for intra-channel fiber nonlinearity compensation in long-haul optical communication systems." *Optics express*, vol. 18, no. 16, pp. 17 075–88, Aug. 2010.

[43] L. Liu, L. Li, Y. Huang, K. Cui, Q. Xiong, F. N. Hauske, C. Xie, and Y. Cai, "Intrachannel Nonlinearity Compensation by Inverse Volterra Series Transfer Function," *Journal of Lightwave Technology*, vol. 30, no. 3, pp. 310–316, Feb. 2012.

[44] X. Liu, A. R. Chraplyvy, P. J. Winzer, R. W. Tkach, and S. Chandrasekhar, "Phase-conjugated twin waves for communication beyond the Kerr nonlinearity limit," *Nature Photonics*, vol. 7, no. 7, pp. 560–568, May 2013.

[45] M. Morshed, A. J. Lowery, and L. B. Du, "Improving performance of optical phase conjugation by splitting the nonlinear element." *Optics express*, vol. 21, no. 4, pp. 4567–77, Feb. 2013.

[46] T. Kobayashi, A. Sano, A. Matsuura, Y. Miyamoto, and K. Ishihara, "Nonlinear Tolerant Spectrally-Efficient Transmission Using PDM 64-QAM Single Carrier FDM With Digital Pilot-Tone," *Journal of Lightwave Technology*, vol. 30, no. 24, pp. 3805–3815, Dec. 2012.

[47] M. I. Yousefi and F. R. Kschischang, "Information Transmission Using the Nonlinear Fourier Transform, Parts I-III," *IEEE Transactions on Information Theory*, vol. 60, no. 7, pp. 4312–4369, Jul. 2014.

[48] C. Lin, S. Chandrasekhar, and P. J. Winzer, "Experimental Study of the Limits of Digital Nonlinearity Compensation in DWDM Systems," in *Optical Fiber Communication Conference*, Mar. 2015, p. Th4D.4.

[49] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "On shaping gain in the nonlinear fiber-optic channel," in *2014 IEEE International Symposium on Information Theory*. IEEE, Jun. 2014, pp. 2794–2798.

[50] A. D. Shiner, M. Reimer, A. Borowiec, S. O. Gharan, J. Gaudette, P. Mehta, D. Charlton, K. Roberts, and M. O'Sullivan, "Demonstration of an 8-dimensional modulation format with reduced inter-channel nonlinearities in a polarization multiplexed coherent system." *Optics express*, vol. 22, no. 17, pp. 20 366–74, Aug. 2014.

[51] I. Djordjevic, W. Ryan, and B. Vasic, *Coding for Optical Channels*. Springer US, 2010.

[52] H. Zhang and H. G. Batshon, "Coded modulation and approaching nonlinear Shannon limit," in *2014 The European Conference on Optical Communication (ECOC)*. IEEE, Sep. 2014, pp. 1–3.

[53] T. Cover and J. Thomas, *Elements of information theory*, ser. Wiley series in telecommunications. Wiley, 1991.

[54] D. J. Costello and G. D. Forney, "Channel coding: The road to channel capacity," *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1150–1177, Jun. 2007. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4282117

[55] E. Agrell, "The channel capacity increases with power," pp. 1–8, Aug. 2011.

[56] E. Agrell, A. Alvarado, G. Durisi, and M. Karlsson, "Capacity of a Nonlinear Optical Channel With Finite Memory," *Journal of Lightwave Technology*, vol. 32, no. 16, pp. 2862–2876, Aug. 2014.

[57] G. Kramer, M. I. Yousefi, and F. R. Kschischang, "Upper Bound on the Capacity of a Cascade of Nonlinear and Noisy Channels," Mar. 2015.

[58] L. Duang, B. Rimoldi, and R. Urbanke, "Approaching the AWGN channel capacity without active shaping," in *Proceedings of IEEE International Symposium on Information Theory.* IEEE, 1997, p. 374.

[59] X. Ma and L. Ping, "Coded Modulation Using Superimposed Binary Codes," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3331–3343, Dec. 2004.

[60] H. S. Cronie, "Superposition Coding for Power-and Bandwidth Efficient Communication over the Gaussian Channel," in *2007 IEEE International Symposium on Information Theory.* IEEE, Jun. 2007, pp. 2311–2315.

[61] J. Tong and L. Ping, "Performance analysis of superposition coded modulation," *Physical Communication*, vol. 3, no. 3, pp. 147–155, Sep. 2010.

[62] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, Jul. 1948.

[63] B. P. Smith and F. R. Kschischang, "A Pragmatic Coded Modulation Scheme for High-Spectral-Efficiency Fiber-Optic Communications," *Journal of Lightwave Technology*, vol. 30, no. 13, pp. 2047–2053, Jul. 2012.

[64] L. Beygi, E. Agrell, J. M. Kahn, and M. Karlsson, "Rate-Adaptive Coded Modulation for Fiber-Optic Communications," *Journal of Lightwave Technology*, vol. 32, no. 2, pp. 333–343, Jan. 2014.

[65] T. Fehenberger, G. Böcherer, A. Alvarado, and N. Hanik, "LDPC coded modulation with probabilistic shaping for optical fiber systems," 2015, pp. 1–3.

[66] G. Böcherer, P. Schulte, and F. Steiner, "Bandwidth Efficient and Rate-Matched Low-Density Parity-Check Coded Modulation," p. 13, feb 2015. [Online]. Available: http://arxiv.org/abs/1502.02733

[67] M. P. Yankov, D. Zibar, K. J. Larsen, L. P. B. Christensen, and S. Forchhammer, "Constellation Shaping for Fiber-Optic Channels With QAM and High Spectral Efficiency," *IEEE Photonics Technology Letters*, vol. 26, no. 23, pp. 2407–2410, Dec. 2014.

[68] F. Buchali, G. Böcherer, W. Idler, L. Schmalen, P. Schulte, and F. Steiner, "Experimental Demonstration of Capacity Increase and Rate-Adaptation by Probabilistically Shaped 64-QAM," in *41st European Conference and Exposition on Optical Communications*. IEEE, sep 2015, p. PDP.3.4.

[69] I. Djordjevic, Tao Liu, Lei Xu, and Ting Wang, "Optimum signal constellation design for high-speed optical transmission," pp. 1–3, 2012.

[70] H. G. Batshon and I. B. Djordjevic, "Iterative Polar Quantization-Based Modulation to Achieve Channel Capacity in Ultrahigh-Speed Optical Communication Systems," *IEEE Photonics Journal*, vol. 2, no. 4, pp. 593–599, Aug. 2010.

[71] T. H. Lotz, X. Liu, S. Chandrasekhar, P. J. Winzer, H. Haunstein, S. Randel, S. Corteselli, B. Zhu, and D. W. Peckham, "Coded PDM-OFDM Transmission With Shaped 256-Iterative-Polar-Modulation Achieving 11.15-b/s/Hz Intrachannel Spectral Efficiency and 800-km Reach," *Journal of Lightwave Technology*, vol. 31, no. 4, pp. 538–545, Feb. 2013.

[72] S. Chandrasekhar, X. Liu, P. J. Winzer, T. H. Lotz, C. J. Youn, Y. H. Kwon, and E. S. Nam, "Demonstration of 3×341-Gb/s PDM-OFDM-256 Iterative Polar Modulation Signals Over 495 km of a Field-Deployed WDM System," *Journal of Lightwave Technology*, vol. 33, no. 3, pp. 625–630, Feb. 2015.

[73] J. Estaran, A. Caballero, D. Zibar, C. Peucheret, and I. Monroy, "Experimental Demonstration of Capacity-Achieving Phase-Shifted Superposition Modulation," in *39th European Conference and Exhibition on Optical Communication (ECOC 2013)*. Institution of Engineering and Technology, 2013, pp. 636–638.

[74] J. Estaran, D. Zibar, and I. T. Monroy, "Capacity-Approaching Superposition Coding for Optical Fiber Links," *Journal of Lightwave Technology*, vol. 32, no. 17, pp. 2960–2972, Sep. 2014.

[75] ——, "Capacity and shaping in coherent fiber-optic links [Invited]," in *2014 IEEE Photonics Conference*. IEEE, Dec. 2014, pp. 306–307.

[76] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[77] J. Ritcey, "Bit-interleaved coded modulation with iterative decoding," in *1999 IEEE International Conference on Communications (Cat. No. 99CH36311)*, vol. 2. IEEE, 1999, pp. 858–863.

[78] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *Proceedings IEEE International Conference on Communications ICC '95*, vol. 2. IEEE, 1995, pp. 1009–1013.

[79] B. Vasic, I. Djordjevic, and V. Rao, "Suppression of intrachannel nonlinearities using BCJR algorithm and iterative decoding," in *2006 Optical Fiber Communication Conference and the National Fiber Optic Engineers Conference*. IEEE, 2006, p. 3 pp.

[80] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (Corresp.)," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.

[81] C. Bishop, *Pattern Recognition and Machine Learning*, ser. Information Science and Statistics. Springer, 2006.

[82] D. Zibar, O. Winther, N. Franceschi, R. Borkowski, A. Caballero, V. Arlunno, M. N. Schmidt, N. G. Gonzales, B. Mao, Y. Ye, K. J. Larsen, and I. T. Monroy, "Nonlinear impairment compensation using expectation maximization for dispersion managed and unmanaged PDM 16-QAM transmission." *Optics express*, vol. 20, no. 26, pp. B181–96, Dec. 2012.

[83] S. T. Brink, "Designing Iterative Decoding Schemes with the Extrinsic Information Transfer Chart," *AEÜ - International Journal of*

Electronics and Communications, vol. 54, no. 6, pp. 389–398, Sep. 2000.

[84] K. Kikuchi, "Polarization-demultiplexing algorithm in the digital coherent receiver," in *2008 Digest of the IEEE/LEOS Summer Topical Meetings*. IEEE, 2008, pp. 101–102.

[85] S. J. Savory, "Digital Coherent Optical Receivers: Algorithms and Subsystems," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 16, no. 5, pp. 1164–1179, sep 2010.

[86] I. Fatadin, D. Ives, and S. Savory, "Blind Equalization and Carrier Phase Recovery in a 16-QAM Optical Coherent System," *Journal of Lightwave Technology*, vol. 27, no. 15, pp. 3042–3049, aug 2009.

[87] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers, Synchronization, Channel Estimation, and Signal Processing*, ser. Digital Communication Receivers. Wiley, 1997.

[88] J. B. Anderson and F. Rusek, "The shannon bit error limit for linear coded modulation," in *Proc., 2004 Intern. Symp. on Information Theory and its Applications, Parma*, 2004, pp. 9–11.

[89] I. Djordjevic, L. Minkov, and H. Batshon, "Mitigation of linear and nonlinear impairments in high-speed optical networks by using LDPC-coded turbo equalization," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 6, pp. 73–83, Aug. 2008.

[90] J. Estaran, M. A. Mestre, P. Jenneve, H. Mardoyan, I. Tafur Monroy, D. Zibar, and S. Bigo, "Coherent Optical Orthogonal Frequency-Division Multiplexing for Optical Slot Switched Intra-Datacenters Networks," in *2015 The European Conference on Optical Communication (ECOC)*. IEEE, Sep. 2015, pp. 1–3.

[91] "OIF Carrier WG Requirements for Intermediate Reach 100G DWDM for Metro Type Applications," Mar. 2014. [Online]. Available: http://www.oiforum.com/public/documents/OIF%20Carrier%20WG%20Requirements%20for%20Intermediate%20Reach%20100G%20DWDM%20for%20Metro%20Type%20Applications.pdf

[92] J. J. G. Torres, A. M. C. Soto, and N. G. Gonzalez, "Characterization of interchannel interference effects in multicarrier 32-Gbaud QPSK/16QAM Nyquist systems," in *IEEE Colombian Conference*

on Communication and Computing (IEEE COLCOM 2015). IEEE, May 2015, pp. 1–6.

[93] X. Liu and S. Chandrasekhar, "Superchannel for Next-Generation Optical Networks," in *Optical Fiber Communication Conference*, Mar. 2014, p. W1H.5.

[94] S. Chandrasekhar, "OFDM Based Superchannel Transmission Technology," *Journal of Lightwave Technology*, vol. 30, no. 24, pp. 3816–3823, Dec. 2012.

[95] J. Armstrong, "OFDM for Optical Communications," *Journal of Lightwave Technology*, vol. 27, no. 3, pp. 189–204, Feb. 2009.

[96] M. I. Olmedo, T. Zuo, J. B. Jensen, Q. Zhong, X. Xu, S. Popov, and I. T. Monroy, "Multiband Carrierless Amplitude Phase Modulation for High Capacity Optical Data Links," *Journal of Lightwave Technology*, vol. 32, no. 4, pp. 798–804, Feb. 2014.

[97] A. Rezania, J. H. Ke, Y. Gao, and J. C. Cartledge, "Single-Carrier 448 Gb/s Dual-Polarization 16-QAM Transmission over 1200 km Using Fixed Look-Up Table Based MAP Detection," in *Optical Fiber Communication Conference*, Mar. 2014, p. Th3E.4.

[98] J. H. Ke, Y. Gao, and J. C. Cartledge, "400 Gbit/s single-carrier and 1 Tbit/s three-carrier superchannel signals using dual polarization 16-QAM with look-up table correction and optical pulse shaping." *Optics express*, vol. 22, no. 1, pp. 71–83, Jan. 2014.

[99] A. Rezania and J. Cartledge, "Transmission Performance of 448 Gb/s Single-Carrier and 1.2 Tb/s Three-Carrier Superchannel Using Dual-Polarization 16-QAM with Fixed LUT Based MAP Detection," *Journal of Lightwave Technology*, vol. PP, no. 99, 2015.

[100] P. J. Winzer, A. H. Gnauck, S. Chandrasekhar, S. Draving, J. Evangelista, and B. Zhu, "Generation and 1,200-km transmission of 448-Gb/s ETDM 56-Gbaud PDM 16-QAM using a single I/Q modulator," in *36th European Conference and Exhibition on Optical Communication*. IEEE, Sep. 2010, pp. 1–3.

[101] Y.-K. Huang, E. Ip, M. Huang, B. Zhu, P. Ji, Y. Shao, D. Peckham, R. Lingle, Y. Aono, T. Tajima, and T. Wang, "10×456-Gb/s

DP-16QAM transmission over 8×100 km of ULAF using coherent detection with a 30-GHz analog-to-digital converter," pp. 1–2, 2010.

[102] J. K. Fischer, L. Molle, M. Nolle, C. Schmidt-Langhorst, J. Hilt, R. Ludwig, D. W. Peckham, and C. Schubert, "8×448-Gb/s WDM Transmission of 56-GBd PDM 16-QAM OTDM Signals Over 250-km Ultralarge Effective Area Fiber," *IEEE Photonics Technology Letters*, vol. 23, no. 4, pp. 239–241, Feb. 2011.

[103] F. Buchali, A. Klekamp, L. Schmalen, and D. Tomislav, "Implementation of 64QAM at 42.66 GBaud Using 1.5 Samples per Symbol DAC and Demonstration of up to 300 km Fiber Transmission," in *Optical Fiber Communication Conference*, 2014, p. M2A.1.

[104] R. Rios-Müller, J. Renaudier, P. Brindel, H. Mardoyan, P. Jennevé, L. Schmalen, and G. Charlet, "1-Terabit/s Net Data-Rate Transceiver Based on Single-Carrier Nyquist-Shaped 124 GBaud PDM-32QAM," in *Optical Fiber Communication Conference Post Deadline Papers*, Mar. 2015, p. Th5B.1.

[105] ——, "Experimental Comparison between Super-channel and Subband Single-Carrier for 400 Gb/s and 800 Gb/s Transport," in *2015 The European Conference on Optical Communication (ECOC)*. IEEE, Sep. 2015, p. Tu.1.4.4.

[106] S. Randel, D. Pilori, S. Corteselli, G. Raybon, A. Adamiecki, A. Gnauck, S. Chandrasekhar, P. J. Winzer, L. Altenhain, A. Bielik, and R. Schmid, "All-Electronic Flexibly Programmable 864-Gb/s Single-Carrier PDM-64-QAM," in *Optical Fiber Communication Conference: Postdeadline Papers*, Mar. 2014, p. Th5C.8.

[107] O. Bertran-Pardo, J. Renaudier, H. Mardoyan, P. Tran, R. Rios-Muller, A. Konczykowska, J.-Y. Dupuy, F. Jorge, M. Riet, B. Duval, J. Godin, S. Randel, G. Charlet, and S. Bigo, "Transmission of 50-GHz-Spaced Single-Carrier Channels at 516Gb/s over 600km," in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013*, Mar. 2013, p. OTh4E.2.

[108] D. Chang, F. Yu, Z. Xiao, Y. Li, N. Stojanovic, C. Xie, X. Shi, X. Xu, and Q. Xiong, "FPGA Verification of a Single QC-LDPC Code for 100 Gb/s Optical Systems without Error Floor down to BER of

10^-15," in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2011*, Mar. 2011, p. OTuN2.

[109] D. Otuya, K. Kasai, T. Hirooka, and M. Nakazawa, "Single-Channel 1.92 Tbit/s, 64 QAM Coherent Nyquist Orthogonal TDM Transmission with a Spectral Efficiency of 10.6 bit/s/Hz," *Journal of Lightwave Technology*, vol. PP, no. 99, pp. 1–1, 2015.

[110] L. H. H. Carvalho, C. Floridia, C. Franciscangelis, V. Parahyba, E. P. da Silva, N. G. Gonzalez, and J. Oliveira, "WDM Transmission of 3x1.12-Tb/s PDM-16QAM Superchannels with 6.5-b/s/Hz in a 162.5-GHz Flexible-Grid using only Optical Spectral Shaping," in *Optical Fiber Communication Conference*, Mar. 2014, p. M3C.3.

[111] Y. Loussouarn, E. Pincemin, M. Song, S. Gauthier, Y. Chen, and Z. Shengqian, "400 Gbps Real-Time Coherent Nyquist-WDM DP-16QAM Transmission over Legacy G.652 or G.655 Fibre Infrastructure with 2 dB Margins," in *Optical Fiber Communication Conference*, Mar. 2015, p. W3E.3.

[112] O. Vassilieva, T. Yamauchi, S. Oda, I. Kim, T. Hoshida, Y. Aoki, J. C. Rasmussen, and M. Sekiya, "Flexible Grid Network Optimization for Maximum Spectral Efficiency and Reach," in *2015 The European Conference on Optical Communication (ECOC)*. IEEE, Sep. 2015, p. Tu.1.4.2.

[113] E. Porto da Silva, L. H. H. Carvalho, J. C. M. Diniz, J. R. Oliveira, V. B. Ribeiro, R. Silva, J. P. K. Perin, M. L. Silva, P. P. G. Cardoso, and J. Oliveira, "448 Gb/s Dual-Carrier PDM-RZ-16QAM on 75-GHz Grid over 720 km with 10 Flexi-Grid ROADM passes," in *Latin America Optics and Photonics Conference*, Nov. 2012, p. LM1C.3.

[114] E. Porto da Silva, L. H. H. Carvalho, C. Franciscangelis, J. Diniz, J. Oliveira, and A. Bordonalli, "Spectrally-Efficient 448-Gb/s dual-carrier PDM-16QAM channel in a 75-GHz grid," in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013*, Mar. 2013, p. JTh2A.39.

[115] M.-F. Huang, S. Zhang, K. Mino, and Y. Aono, "Transmission of 400G Dual-Carrier DP-16QAM and Multi-Carrier DP-QPSK Signals over Regional and Long-Haul Distances with Span Lengths Greater

than 200 km," in *Optical Fiber Communication Conference*, Mar. 2014, p. Th4F.3.

[116] K. Sugihara, Y. Miyata, T. Sugihara, K. Kubo, H. Yoshida, W. Matsumoto, and T. Mizuochi, "A Spatially-coupled Type LDPC Code with an NCG of 12 dB for Optical Transmission beyond 100 Gb/s," in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013*, Mar. 2013, p. OM2B.4.

[117] J. Renaudier, R. R. Muller, L. Schmalen, P. Tran, P. Brindel, and G. Charlet, "1-Tb/s PDM-32QAM superchannel transmission at 6.7-b/s/Hz over SSMF and 150-GHz-grid ROADMs," in *2014 The European Conference on Optical Communication (ECOC)*. IEEE, Sep. 2014.

[118] X. Zhou, L. E. Nelson, P. Magill, R. Isaac, B. Zhu, D. W. Peckham, P. I. Borel, and K. Carlson, "PDM-Nyquist-32QAM for 450-Gb/s Per-Channel WDM Transmission on the 50 GHz ITU-T Grid," *Journal of Lightwave Technology*, vol. 30, no. 4, pp. 553–559, Feb. 2012.

[119] F. Chang, K. Onohara, and T. Mizuochi, "Forward error correction for 100 G transport networks," *IEEE Communications Magazine*, vol. 48, no. 3, pp. S48–S55, Mar. 2010.

[120] X. Zhou, L. Nelson, R. Issac, P. Magill, B. Zhu, and D. Peckham, "1200km Transmission of 50GHz spaced, 5x504-Gb/s PDM-32-64 hybrid QAM using Electrical and Optical Spectral Shaping," in *Optical Fiber Communication Conference*, Mar. 2012, p. OM2A.2.

[121] T. Kobayashi, A. Sano, A. Matsuura, M. Yoshida, T. Sakano, H. Kubota, Y. Miyamoto, K. Ishihara, M. Mizoguchi, and M. Nagatani, "45.2Tb/s C-band WDM transmission over 240km using 538Gb/s PDM-64QAM single carrier FDM signal with digital pilot tone," in *37th European Conference and Exposition on Optical Communications*, Sep. 2011, p. Th.13.C.6.

[122] D. Qian, M.-F. Huang, E. Ip, Y.-K. Huang, Y. Shao, J. Hu, and T. Wang, "High Capacity/Spectral Efficiency 101.7-Tb/s WDM Transmission Using PDM-128QAM-OFDM Over 165-km SSMF Within C- and L-Bands," *Journal of Lightwave Technology*, vol. 30, no. 10, pp. 1540–1548, May 2012.

[123] Chao Li, Xuebing Zhang, Haibo Li, Cai Li, Min Lou, Zhaohui Li, Jing Xu, Qi Yang, and Shaohua Yu, "Experimental Demonstration of 429.96-Gb/s OFDM/OQAM–64QAM Over 400-km SSMF Transmission Within a 50-GHz Grid," *IEEE Photonics Journal*, vol. 6, no. 4, pp. 1–8, Aug. 2014.

[124] A. Pagano, E. Riccardi, M. Bertolini, V. Farelli, and T. Van De Velde, "400Gb/s Real-time Trial Using Rate-adaptive Transponders for Next Generation Flexible-grid Networks," in *Optical Fiber Communication Conference*, Mar. 2014, p. Tu2B.4.

[125] B. Lavigne, "400Gb/s Trials on Commercial Systems Using Real-time Bit-rate-adaptive Transponders for Next Generation Networks," in *Optical Fiber Communication Conference*, Mar. 2015, p. W3E.1.

[126] M. Alfiad and S. Tibuleac, "100G Superchannel Transmission Using 4×28 Gb/s Subcarriers on a 25-GHz Grid," *IEEE Photonics Technology Letters*, vol. 27, no. 2, pp. 157–160, Jan. 2015.

[127] J. Estaran, M. Iglesias, D. Zibar, X. Xu, and I. Tafur Monroy, "First Experimental Demonstration of Coherent CAP for 300-Gb/s Metropolitan Optical Networks," in *Optical Fiber Communication Conference*, Mar. 2014, p. Th3K.3.

[128] D. D. Falconer, "Carrierless AM/PM," Bell Laboratories, Tech. Rep., Jul. 1975.

[129] J.-J. Werner, "Tutorial on carrierless AM/PM - Part I and II," ANSI X3T9. 5TP/PMD Working Group, Tech. Rep., 1993.

[130] J. Proakis, *Digital Communications*, ser. McGraw-Hill Series in Electrical and Computer Engineering. Computer Engineering. McGraw-Hill, 2001.

[131] D. S. Millar, T. Koike-Akino, S. O. Arık, K. Kojima, K. Parsons, T. Yoshida, and T. Sugihara, "High-dimensional modulation for coherent optical communications systems." *Optics express*, vol. 22, no. 7, pp. 8798–812, Apr. 2014.

[132] S. Ishimura and K. Kikuchi, "Multi-dimensional permutation-modulation format for coherent optical communications." *Optics express*, vol. 23, no. 12, pp. 15 587–97, Jun. 2015.

[133] M. I. Olmedo, A. Tatarczak, T. Zuo, J. Estaran, X. Xu, and I. Tafur Monroy, "Towards 100 Gbps over 100m MMF using a 850nm VCSEL," in *Optical Fiber Communication Conference*, 2014, p. M2E.5.

[134] T. Zuo, A. Tatarczak, M. Iglesias, J. Estaran, J. Bevensee Jensen, Q. Zhong, X. Xu, and I. Tafur, "O-band 400 Gbit/s Client Side Optical Transmission Link," in *Optical Fiber Communication Conference*, 2014, p. M2E.4.

[135] A. Tatarczak, M. I. Olmedo, T. Zuo, J. Estaran, J. Bevensee Jensen, X. Xu, and I. tafur Monroy, "Enabling 4-Lane Based 400G Client-Side Transmission Links with MultiCAP Modulation," *Advances in Optical Technologies*, vol. 2015, 2015.

[136] Li Xinrong, I. Thng, and Ko Chi Chung, "An improved design of 3-D CAP signature waveforms against quantization noise," in *Proceedings IEEE International Symposium on Computers and Communications (Cat. No.PR00250)*. IEEE Comput. Soc, 1999, pp. 369–374.

[137] A. Shalash and K. Parhi, "Multidimensional carrierless AM/PM systems for digital subscriber loops," *IEEE Transactions on Communications*, vol. 47, no. 11, pp. 1655–1667, 1999.

[138] I. Thng, "A new 3D CAP system," in *Proceedings of IEEE. IEEE Region 10 Conference. TENCON 99. 'Multimedia Technology for Asia-Pacific Information Infrastructure' (Cat. No.99CH37030)*, vol. 1. IEEE, 1999, pp. 309–312.

[139] ——, "A new digital approach to design 3-d cap waveforms," *IEEE Transactions on Communications*, vol. 51, no. 1, pp. 12–16, Jan. 2003.

[140] G. Stepniak, "Comparison of Efficiency of N-Dimensional CAP Modulations," *Journal of Lightwave Technology*, vol. 32, no. 14, pp. 2516–2523, Jul. 2014.

[141] X. Tang and I.-J. Thng, "An NS Frequency-Domain Approach for Continuous-Time Design of CAP/ICOM Waveform," *IEEE Transactions on Communications*, vol. 52, no. 12, pp. 2154–2164, dec 2004.

[142] S. Daumont, B. Rihawi, and Y. Lout, "Root-Raised Cosine filter influences on PAPR distribution of single carrier signals," in *2008*

3rd International Symposium on Communications, Control and Signal Processing. IEEE, mar 2008, pp. 841–845.

[143] Tao Jiang and Yiyan Wu, "An Overview: Peak-to-Average Power Ratio Reduction Techniques for OFDM Signals," *IEEE Transactions on Broadcasting*, vol. 54, no. 2, pp. 257–268, jun 2008.

[144] S. Thompson, J. Proakis, and J. Zeidler, "The effectiveness of signal clipping for PAPR and total degradation reduction in OFDM systems," in *GLOBECOM '05. IEEE Global Telecommunications Conference, 2005.*, vol. 5. IEEE, 2005, pp. 5 pp.–2811.

[145] M. B. Othman, M. Wieckowski, J. B. Jensen, and I. T. Monroy, "Experimental Investigations of 3-D-/4-D-CAP Modulation With Directly Modulated VCSELs," *IEEE Photonics Technology Letters*, vol. 24, no. 22, pp. 2009–2012, nov 2012.

[146] S. Loquai, R. Kruglov, B. Schmauss, C.-A. Bunge, F. Winkler, O. Ziemann, E. Hartl, and T. Kupfer, "Comparison of Modulation Schemes for 10.7 Gb/s Transmission Over Large-Core 1 mm PMMA Polymer Optical Fiber," *Journal of Lightwave Technology*, vol. 31, no. 13, pp. 2170–2176, jul 2013.

[147] G. Stepniak, L. Maksymiuk, and J. Siuzdak, "Experimental Comparison of PAM, CAP, and DMT Modulations in Phosphorescent White LED Transmission Link," *IEEE Photonics Journal*, vol. 7, no. 3, pp. 1–8, jun 2015.

[148] Y. Wang, L. Tao, Y. Wang, and N. Chi, "High Speed WDM VLC System Based on Multi-Band CAP64 With Weighted Pre-Equalization and Modified CMMA Based Post-Equalization," *IEEE Communications Letters*, vol. 18, no. 10, pp. 1719–1722, Oct. 2014.

[149] T. Xu, G. Jacobsen, S. Popov, J. Li, E. Vanin, K. Wang, A. T. Friberg, and Y. Zhang, "Chromatic dispersion compensation in coherent transmission system using digital filters." *Optics express*, vol. 18, no. 15, pp. 16 243–57, jul 2010.

[150] R. Borkowski, D. Zibar, and I. Tafur Monroy, "Anatomy of a Digital Coherent Receiver," *IEICE Transactions on Communications*, vol. E97.B, no. 8, pp. 1528–1536, aug 2014.

[151] S. Kumar, *Impact of Nonlinearities on Fiber Optic Communications*, ser. Optical and Fiber Communications Reports. Springer New York, 2011.

[152] G. Yi, L. Gang, and G. Jianhua, "A novel time and frequency synchronization scheme for OFDM systems," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 2, pp. 321–325, may 2008.

[153] T. Schmidl and D. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Transactions on Communications*, vol. 45, no. 12, pp. 1613–1621, 1997.

[154] X. Zhang, X. Pang, L. Deng, D. Zibar, I. T. Monroy, and R. Younce, "High phase noise tolerant pilot-tone-aided DP-QPSK optical communication systems." *Optics express*, vol. 20, no. 18, pp. 19 990–5, aug 2012.

[155] G. Jacobsen, T. Xu, S. Popov, J. Li, A. T. Friberg, and Y. Zhang, "Receiver implemented RF pilot tone phase noise mitigation in coherent optical nPSK and nQAM systems." *Optics express*, vol. 19, no. 15, pp. 14 487–94, jul 2011.

[156] L. B. Du and A. J. Lowery, "Pilot-based cross-phase modulation compensation for coherent optical orthogonal frequency division multiplexing long-haul optical communications systems." *Optics letters*, vol. 36, no. 9, pp. 1647–9, may 2011.

[157] S. Randel, S. Adhikari, and S. L. Jansen, "Analysis of RF-Pilot-Based Phase Noise Compensation for Coherent Optical OFDM Systems," *IEEE Photonics Technology Letters*, vol. 22, no. 17, pp. 1288–1290, sep 2010.

[158] S. Jansen, I. Morita, T. Schenk, N. Takeda, and H. Tanaka, "Coherent Optical 25.8-Gb/s OFDM Transmission Over 4160-km SSMF," *Journal of Lightwave Technology*, vol. 26, no. 1, pp. 6–15, jan 2008.

[159] S. Mumtaz, G.-B. Othman, and Y. Jaouen, "Space-Time Codes for Optical Fiber Communication with Polarization Multiplexing," in *2010 IEEE International Conference on Communications*. IEEE, may 2010, pp. 1–5.

[160] M. Kuschnerov, M. Chouayakh, K. Piyawanno, B. Spinnler, E. de Man, P. Kainzmaier, M. S. Alfiad, A. Napoli, and B. Lankl,

"Data-Aided Versus Blind Single-Carrier Coherent Receivers," *IEEE Photonics Journal*, vol. 2, no. 3, pp. 387–403, jun 2010.

[161] F. Pittala, I. Slim, A. Mezghani, and J. A. Nossek, "Training-Aided Frequency-Domain Channel Estimation and Equalization for Single-Carrier Coherent Optical Transmission Systems," *Journal of Lightwave Technology*, vol. 32, no. 24, pp. 4849–4863, dec 2014.

[162] P. Hill, R. Olshansky, and M. Abdollahian, "Novel carrier and clock-recovery circuit for multigigabit/second lightwave systems," *IEEE Photonics Technology Letters*, vol. 5, no. 1, pp. 96–98, jan 1993.

[163] D. Kim, "Use of a pilot tone for clock and carrier recovery in mutlicarrier transmission," Feb. 1 2006, eP Patent App. EP20,010,114,731. [Online]. Available: https://www.google.com/patents/EP1220505A3?cl=en

[164] Jinbiao Xu and Yumin Wang, "New decision-directed equalization algorithm for QAM communication systems," in *Proceedings of GLOBECOM'96. 1996 IEEE Global Telecommunications Conference*, vol. 2. IEEE, 1996, pp. 1330–1334.

[165] D. Zibar, L. H. H. de Carvalho, M. Piels, A. Doberstein, J. Diniz, B. Nebendahl, C. Franciscangelis, J. Estaran, H. Haisch, N. G. Gonzalez, J. C. R. F. de Oliveira, and I. T. Monroy, "Application of Machine Learning Techniques for Amplitude and Phase Noise Characterization," *Journal of Lightwave Technology*, vol. 33, no. 7, pp. 1333–1343, apr 2015.

[166] M. Piels, M. Iglesias Olmedo, X. Pang, R. Schatz, G. Jacobsen, S. Popov, and D. Zibar, "Rate Equation-Based Phase Recovery for Semiconductor Laser Coherent Transmitters," in *Optical Fiber Communication Conference*, mar 2015, p. W1E.7.

[167] B. Li, K. J. Larsen, D. Zibar, and I. Tafur Monroy, "Over 10 dB Net Coding Gain Based on 20% Overhead Hard Decision Forward Error Correction in 100G Optical Communication Systems," in *37th European Conference and Exposition on Optical Communications*, sep 2011, p. Tu.6.A.3.

[168] J. Justesen, "Performance of Product Codes and Related Structures with Iterated Decoding," *IEEE Transactions on Communications*, vol. 59, no. 2, pp. 407–415, feb 2011.

[169] Y. Jiang, "The EGN Model of Nonlinear Propagation in Coherent Optical Transmission Systems and Its Applications," Ph.D. dissertation, Politecnico di Torino, dec 2014.

[170] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "A Simple and Effective Closed-Form GN Model Correction Formula Accounting for Signal Non-Gaussian Distribution," *Journal of Lightwave Technology*, vol. 33, no. 2, pp. 459–473, jan 2015.

[171] "The 2015 Ethernet Roadmap," Ethernet Alliance, Tech. Rep., 2015.

[172] "InfiniBand Roadmap." [Online]. Available: http://www.infinibandta.org/content/pages.php?pg=technology_overview

[173] S. Kipp, "FCIA Official Speedmap v20," Fiber Channel Industry Association, Tech. Rep., 2015.

[174] J. D'Ambrosia and D. Law, "Objective Considerations: Given Today's Data Center Environment," in *Applications Ad Hoc*. IEEE 802.3 400Gb/s Ethernet Study Group, Oct. 2013. [Online]. Available: http://www.ieee802.org/3/400GSG/public/adhoc/app/dambrosia_app_01_1013.pdf

[175] P. Polishuk, "Opportunities for Fiber Optics Less Than 100 Meters," in *FIBERFEST SYMPOSIUM: New Opportunities in Fiber Optics Applications*. IGI Consulting, Inc., 2009. [Online]. Available: http://www.igigroup.net/download/FFS09/0900_Polishuk_IGI.pdf

[176] "Wideband Multimode Fiber – What is it and why does it make sense?" COMMSCOPE, Tech. Rep., Jun. 2015. [Online]. Available: file:///C:/Users/jome/Downloads/Wideband_Multimode_Fiber_What_Why_WP-109042.pdf

[177] J. D'Ambrosia, "IEEE P802.3bs Baseline Summary," Jul. 2015. [Online]. Available: http://www.ieee802.org/3/bs/baseline_3bs_0715.pdf

[178] C. Cole, J. J. Maki, A. Srivastava, and P. Stassar, "400Gb/s 8x50G PAM4 WDM 2km SMF PMD Baseline Specifications," Jul. 2015. [Online]. Available: http://www.ieee802.org/3/bs/public/15_07/cole_3bs_01a_0715.pdf

[179] D. Dove and X. Song, "Considering 2 km and 10 km Architectural Alternatives," Jan. 2014. [Online]. Available: http://www.ieee802.org/3/400GSG/public/14_01/dove_400_01a_0114.pdf

[180] B. Welch, "100 Gbps/λ vs. 50 Gbps/λ : Cost and Power Comparisons in Silicon Photonics," Jan. 2015. [Online]. Available: http://www.ieee802.org/3/bs/public/15_01/welch_3bs_02_0115.pdf

[181] "IEEE P802.3bj 100 Gb/s Backplane and Copper Cable Task Force," Jul. 2014. [Online]. Available: http://www.ieee802.org/3/bj/

[182] M. Langhammer, "400GE FEC Breakout Architecture Analysis," Jul. 2015. [Online]. Available: http://www.ieee802.org/3/bs/public/15_07/langhammer_3bs_02_0715.pdf

[183] X. Song, X. Man, and S. Shen, "Considerations on the Telecom Application for 400GbE," Nov. 2013. [Online]. Available: http://www.ieee802.org/3/400GSG/public/13_11/song_x_400_01a_1113.pdf#page=6

[184] A. Lowery and J. Armstrong, "Orthogonal-frequency-division multiplexing for dispersion compensation of long-haul optical systems," *Optics Express*, vol. 14, no. 6, p. 2079, Mar. 2006.

[185] W.-R. Peng, B. Zhang, K.-M. Feng, X. Wu, A. E. Willner, and S. Chi, "Spectrally Efficient Direct-Detected OFDM Transmission Incorporating a Tunable Frequency Gap and an Iterative Detection Techniques," *Journal of Lightwave Technology*, vol. 27, no. 24, pp. 5723–5735, Dec. 2009.

[186] Zizheng Cao, Jianjun Yu, Wenpei Wang, Lin Chen, and Ze Dong, "Direct-Detection Optical OFDM Transmission System Without Frequency Guard Band," *IEEE Photonics Technology Letters*, vol. 22, no. 11, pp. 736–738, Jun. 2010.

[187] X. Chen, D. Che, A. Li, J. He, and W. Shieh, "Signal-carrier interleaved optical OFDM for direct detection optical communication." *Optics express*, vol. 21, no. 26, pp. 32 501–7, Dec. 2013.

[188] X. Chen, A. Li, D. Che, Q. Hu, Y. Wang, J. He, and W. Shieh, "Block-wise phase switching for double-sideband direct detected optical OFDM signals." *Optics express*, vol. 21, no. 11, pp. 13 436–41, Jun. 2013.

[189] A. Al Amin, H. Takahashi, I. Morita, and H. Tanaka, "100-Gb/s Direct-Detection OFDM Transmission on Independent Polarization Tributaries," *IEEE Photonics Technology Letters*, vol. 22, no. 7, pp. 468–470, Apr. 2010.

[190] D. Qian, N. Cvijetic, Junqiang Hu, and Ting Wang, "108 Gb/s OFDMA-PON With Polarization Multiplexing and Direct Detection," *Journal of Lightwave Technology*, vol. 28, no. 4, pp. 484–493, Feb. 2010.

[191] C.-C. Wei, C.-T. Lin, and C.-Y. Wang, "PMD tolerant direct-detection polarization division multiplexed OFDM systems with MIMO processing." *Optics express*, vol. 20, no. 7, pp. 7316–22, Mar. 2012.

[192] C. Li, H. Li, Q. Yang, M. Luo, X. Zhang, R. Hu, Z. Li, W. Li, and S. Yu, "Single photodiode direct detection system of 100-Gb/s OFDM/OQAM-64QAM over 80-km SSMF within a 50-GHz optical grid." *Optics express*, vol. 22, no. 19, pp. 22 490–7, Sep. 2014.

[193] B. Schmidt, Z. Zan, L. Du, and A. Lowery, "120 Gbit/s Over 500-km Using Single-Band Polarization-Multiplexed Self-Coherent Optical OFDM," *Journal of Lightwave Technology*, vol. 28, no. 4, pp. 328–335, Feb. 2010.

[194] M. Nazarathy and A. Agmon, "Doubling direct-detection data rate by polarization multiplexing of 16-QAM without active polarization control." *Optics express*, vol. 21, no. 26, pp. 31 998–2012, Dec. 2013.

[195] K. Kikuchi, "Electronic polarization-division demultiplexing based on digital signal processing in intensity-modulation direct-detection optical communication systems." *Optics express*, vol. 22, no. 2, pp. 1971–80, Jan. 2014.

[196] C. Brosseau, *Fundamentals of polarized light: a statistical optics approach*, ser. Wiley-interscience publication. John Wiley, 1998.

[197] K. Kikuchi, "Simple and Efficient Algorithm for Polarization Tracking and Demultiplexing in Dual-polarization IM/DD Systems," in *Optical Fiber Communication Conference*. Optical Society of America, Mar. 2015, p. Th1E.3.

[198] M. Morsy-Osman, M. Chagnon, M. Poulin, S. Lessard, and D. V. Plant, "224-Gb/s 10-km Transmission of PDM PAM-4 at 1.3 μm Using a Single Intensity-Modulated Laser and a Direct-Detection MIMO DSP-Based Receiver," *Journal of Lightwave Technology*, vol. 33, no. 7, pp. 1417–1424, Apr. 2015.

[199] M. Chagnon, M. Osman, D. Patel, V. Veerasubramanian, A. Samani, and D. Plant, "1 λ, 6 bits/symbol, 280 and 350 Gb/s Direct Detection Transceiver using Intensity Modulation, Polarization Multiplexing, and Inter-Polarization Phase Modulation," in *Optical Fiber Communication Conference Post Deadline Papers*. Optical Society of America, Mar. 2015, p. Th5B.2.

[200] S. Betti, F. Curti, B. Daino, G. de Marchis, and E. Iannone, "State of polarisation and phase noise independent coherent optical transmission system based on Stokes parameter detection," pp. 1460–1461, 1988.

[201] S. Betti, F. Curti, G. De Marchis, and E. Iannone, "Multilevel coherent optical system based on Stokes parameters modulation," *Journal of Lightwave Technology*, vol. 8, no. 7, pp. 1127–1136, Jul. 1990.

[202] S. Betti, G. De Marchis, and E. Iannone, "Polarization modulated direct detection optical transmission systems," *Journal of Lightwave Technology*, vol. 10, no. 12, pp. 1985–1997, 1992.

[203] S. Benedetto and P. Poggiolini, "Performance evaluation of multilevel polarisation shift keying modulation schemes," *Electronics Letters*, vol. 26, no. 4, p. 244, 1990.

[204] ——, "Theory of polarization shift keying modulation," *IEEE Transactions on Communications*, vol. 40, no. 4, pp. 708–721, Apr. 1992.

[205] ——, "Multilevel polarization shift keying: optimum receiver structure and performance evaluation," *IEEE Transactions on Communications*, vol. 42, no. 2/3/4, pp. 1174–1186, Feb. 1994.

[206] S. Benedetto, A. Djupsjobacka, B. Lagerstrom, R. Paoletti, P. Poggiolini, and G. Mijic, "Multilevel polarization modulation using a specifically designed LiNbO/sub 3/ device," *IEEE Photonics Technology Letters*, vol. 6, no. 8, pp. 949–951, Aug. 1994.

[207] S. Benedetto, R. Gaudino, and P. Poggiolini, "Direct detection of optical digital transmission based on polarization shift keying modulation," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 3, pp. 531–542, Apr. 1995.

[208] M. Nazarathy and E. Simony, "Stokes Space Optimal Detection of Multidifferential Phase and Polarization Shift Keying Modulation," *Journal of Lightwave Technology*, vol. 24, no. 5, p. 1978, May 2006.

[209] ——, "Generalized Stokes parameters-shift keying: a new perspective on optimal detection over electrical and optical vector incoherent channels," *IEEE Transactions on Communications*, vol. 54, no. 3, pp. 499–509, Mar. 2006.

[210] K. Kikuchi and S. Kawakami, "Multi-level signaling in the Stokes space and its application to large-capacity optical communications." *Optics express*, vol. 22, no. 7, pp. 7374–87, Apr. 2014.

[211] ——, "16-ary Stokes-vector Modulation Enabling DSP-based Direct Detection at 100 Gbit/s," in *Optical Fiber Communication Conference*. Optical Society of America, Mar. 2014, p. Th3K.6.

[212] W. Shieh, D. Che, Q. Hu, and A. Li, "Linearization of Optical Channels with Stokes Vector Direct Detection," in *Optical Fiber Communication Conference*. Optical Society of America, Mar. 2015, p. Th1E.5.

[213] D. Che, A. Li, X. Chen, Q. Hu, Y. Wang, and W. Shieh, "160-Gb/s Stokes Vector Direct Detection for Short Reach Optical Communication," in *Optical Fiber Communication Conference: Postdeadline Papers*. Optical Society of America, 2014, p. Th5C.7.

[214] A. Li, D. Che, V. Chen, and W. Shieh, "Spectrally efficient optical transmission based on Stokes vector direct detection." *Optics express*, vol. 22, no. 13, pp. 15 662–7, Jun. 2014.

[215] D. Che, Q. Hu, X. Chen, A. Li, and W. Shieh, "1 - Tb/s Stokes Vector Direct Detection over 480 - km SSMF Transmission," in *Optoelectronics and Communications Conference Post Deadline Papers*, 2014, p. THPDP1.

[216] D. Che, A. Li, X. Chen, Q. Hu, Y. Wang, and W. Shieh, "Stokes vector direct detection for short-reach optical communication." *Optics letters*, vol. 39, no. 11, pp. 3110–3, Jun. 2014.

[217] ——, "Stokes Vector Direct Detection for Linear Complex Optical Channels," *Journal of Lightwave Technology*, vol. 33, no. 3, pp. 678–684, Feb. 2015.

[218] Q. Hu, D. Che, Y. Wang, and W. Shieh, "PMD induced impairment mitigation in Stokes vector direct detection systems," in *Optical Fiber Communication Conference*, Mar. 2015, p. Th1E.2.

[219] Q. Hu, D. Che, and W. Shieh, "Mitigation of PMD induced nonlinear noise in stokes vector direct detection systems," in *2014 The European Conference on Optical Communication (ECOC)*, Sep. 2014, pp. 1–3.

[220] J. Estaran, M. A. Usuga, E. Porto, M. Piels, M. I. Olmedo, and I. T. Monroy, "Quad-polarization transmission for high-capacity IM/DD links [Postdeadline]," in *2014 The European Conference on Optical Communication (ECOC)*. IEEE, Sep. 2014, pp. 1–3.

[221] J. Estaran, M. A. Usuga, E. P. da Silva, M. Piels, M. I. Olmedo, D. Zibar, and I. T. Monroy, "Quaternary Polarization-Multiplexed Subsystem for High-Capacity IM/DD Optical Data Links [Invited]," *Journal of Lightwave Technology*, vol. 33, no. 7, pp. 1408–1416, Apr. 2015.

[222] J. Estaran, X. Lu, D. Zibar, and I. Tafur Monroy, "Stokes Space in Direct-Detection Data Transmission Systems [Invited]," in *2015 Asia Communications and Photonics Conference (ACP)*, Nov. 2015.

[223] C. Herard and A. Lacourt, "Three channel multiplexing using polarization of light," *Optics Communications*, vol. 60, no. 1-2, pp. 27–31, Oct. 1986.

[224] ——, "New multiplexing technique using polarization of light." *Appl. Opt.*, vol. 30, no. 2, pp. 222–31, Jan. 1991.

[225] H. Yaffe and Y. Barad, "Requirements for State of Polarization (SOP) Scrambling in Digital Equalizer and Transponder Verification," Feb. 2011. [Online]. Available: http://www.newridgetech.com/documents/Requirements%20for%20State%20of%20Polarization%20(SOP)%20Scrambling%20in%20EDC%20CDR%20and%20Transponder%20Verification%204.pdf

[226] B. Lu and B. Evans, "Channel equalization by feedforward neural networks," in *ISCAS'99. Proceedings of the 1999 IEEE International*

Symposium on Circuits and Systems VLSI (Cat. No.99CH36349), vol. 5. IEEE, 1999, pp. 587–590.

[227] D. Che, A. Li, Q. Hu, X. Chen, and W. Shieh, "Implementing Simplified Stokes Vector Receiver for Phase Diverse Direct Detection," in *Optical Fiber Communication Conference*, Mar. 2015, p. Th1E.4.

[228] D. Sim, H. Kim, and Y. Chung, "Direct-Detection Receiver for Polarization-Division-Multiplexed OOK Signals," *IEEE Photonics Technology Letters*, vol. PP, no. 99, pp. 1–1, 2015.

[229] I. Ganga, "IEEE P802.3ba Baseline Summary," May 2008. [Online]. Available: http://www.ieee802.org/3/ba/BaselineSummary_0908.pdf

[230] D. L. Peterson, P. J. Leo, and K. B. Rochford, "Field measurements of state of polarization and PMD from a tier-1 carrier," in *Optical Fiber Communication Conference*. Optical Society of America, Feb. 2004, p. FI1.

[231] M. Scholten, T. Coe, J. Dillard, and F. Chang, "Enhanced FEC for 40G / 100G," Sep. 2009. [Online]. Available: http://conference.vde.com/ecoc-2009/programs/documents/ws1_06_mikescholten.pdf

[232] B. Li, K. J. Larsen, J. J. V. Olmos, D. Zibar, and I. T. Monroy, "Application of beyond bound decoding for high speed optical communications," in *Asia Communications and Photonics Conference 2013*. Optical Society of America, 2013, p. AF4C.6.

[233] S. Saldaña Cercós, M. Piels, J. Estarán, M. Usuga, E. Porto da Silva, A. Manolova Fagertun, and I. Tafur Monroy, "100 Gbps IM/DD links using quad-polarization: Performance, complexity, and power dissipation," *Optics Express*, vol. 23, no. 15, p. 19954, Jul. 2015.